PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07K 19/00, C12N 15/62, 15/70, 1/21

A2

(11) International Publication Number: **WO 00/24782**

(43) International Publication Date: 4 May 2000 (04.05.00)

(21) International Application Number: PCT/US99/25044

(22) International Filing Date: 25 October 1999 (25.10.99)

(30) Priority Data:

60/105,371	23 October 1998 (23.10.98)	US
09/428,082	22 October 1999 (22.10.99)	US

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(71) Applicant: AMGEN INC. [US/US]; One Amgen Center Drive, Thousand Oaks, CA 91320–1799 (US).

(72) Inventors: FEIGE, Ulrich; 3117 Deer Valley Avenue, Newbury Park, CA 91320 (US). LIU, Chuan–Fa; 1425 Clover Creek Drive, Longmont, CO 80503 (US). CHEETHAM, Janet; 1695 East Valley Road, Montecito, CA 93108 (US). BOONE, Thomas, Charles; 3010 Deer Valley Avenue, Newbury Park, CA 91320 (US).

(74) Agents: ODRE, Steven, M. et al.; Amgen, Inc., One Amgen Center Drive, Thousand Oaks, CA 91320–1799 (US).

(54) Title: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS

(57) Abstract

The present invention concerns fusion of Fc domains with biologically active peptides and a process for preparing pharmaceutical agents using biologically active peptides. In this invention, pharmacologically active compounds are prepared by a process comprising: a) selecting at least one peptide that modulates the activity of a protein of interest; and b) preparing a pharmacologic agent comprising an Fc domain covalently linked to at least one amino acid of the selected peptide. Linkage to the vehicle increases the half–life of the peptide, which otherwise would be quickly degraded *in vivo*. The preferred vehicle is an Fc domain. The peptide is preferably selected by phage display, *E. coli* display, ribosome display, RNA–peptide screening, or chemical–peptide screening.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Modified Peptides as Therapeutic Agents

Background of the Invention

Recombinant proteins are an emerging class of therapeutic agents. Such recombinant therapeutics have engendered advances in protein formulation and chemical modification. Such modifications can protect therapeutic proteins, primarily by blocking their exposure to proteolytic enzymes. Protein modifications may also increase the therapeutic protein's stability, circulation time, and biological activity. A review article describing protein modification and fusion proteins is Francis (1992), Focus on Growth Factors 3:4-10 (Mediscript, London), which is hereby incorporated by reference.

One useful modification is combination with the "Fc" domain of an antibody. Antibodies comprise two functionally independent parts, a variable domain known as "Fab", which binds antigen, and a constant domain known as "Fc", which links to such effector functions as complement activation and attack by phagocytic cells. An Fc has a long serum half-life, whereas an Fab is short-lived. Capon et al. (1989), Nature 337: 525-31. When constructed together with a therapeutic protein, an Fc domain can provide longer half-life or incorporate such functions as Fc receptor binding, protein A binding, complement fixation and perhaps even placental transfer. Id. Table 1 summarizes use of Fc fusions known in the art.

Table 1—Fc fusion with therapeutic proteins

Form of Fc	Fusion partner	Therapeutic implications	Reference
IgG1	N-terminus of CD30-L	Hodgkin's disease; anaplastic lymphoma; T-cell leukemia	U.S. Patent No. 5,480,981
Murine Fcγ2a	IL-10	anti-inflammatory; transplant rejection	Zheng et al. (1995), J. Immunol. 154: 5590-600
IgG1	TNF receptor	septic shock	Fisher et al. (1996), N. Engl. J. Med. 334: 1697-1702; Van Zee, K. et al. (1996), J. Immunol. 156: 2221-30
IgG, IgA, IgM, or IgE (excluding the first domain)	TNF receptor	inflammation, autoimmune disorders	U.S. Pat. No. 5,808,029, issued September 15, 1998
IgG1	CD4 receptor	AIDS	Capon et al. (1989), Nature 337: 525-31
IgG1, IgG3	N-terminus of IL-2	anti-cancer, antiviral	Harvill et al. (1995), Immunotech. 1: 95-105
IgG1	C-terminus of OPG	osteoarthritis; bone density	WO 97/23614, published July 3, 1997
IgG1	N-terminus of leptin	anti-obesity	PCT/US 97/23183, filed December 11, 1997
Human Ig Cγ1	CTLA-4	autoimmune disorders	Linsley (1991), J. Exp. Med. 174:561-9

A much different approach to development of therapeutic agents is peptide library screening. The interaction of a protein ligand with its receptor often takes place at a relatively large interface. However, as demonstrated for human growth hormone and its receptor, only a few key residues at the interface contribute to most of the binding energy. Clackson et al. (1995), Science 267: 383-6. The bulk of the protein ligand merely displays the binding epitopes in the right topology or serves functions unrelated to binding. Thus, molecules of only "peptide" length (2 to 40 amino acids) can bind to the receptor protein of a given large protein ligand. Such peptides may mimic the bioactivity of the large protein ligand ("peptide agonists") or, through competitive binding, inhibit the bioactivity of the large protein ligand ("peptide antagonists").

Phage display peptide libraries have emerged as a powerful method in identifying such peptide agonists and antagonists. See, for example, Scott et al. (1990), Science 249: 386; Devlin et al. (1990), Science 249: 404; U.S. Pat. No. 5,223,409, issued June 29, 1993; U.S. Pat. No. 5,733,731, issued March 31, 1998; U.S. Pat. No. 5,498,530, issued March 12, 1996; U.S. Pat. No. 5,432,018, issued July 11, 1995; U.S. Pat. No. 5,338,665, issued August 16, 1994; U.S. Pat. No. 5,922,545, issued July 13, 1999; WO 96/40987, published December 19, 1996; and WO 98/15833, published April 16, 1998 (each of which is incorporated by reference). In such libraries, random peptide sequences are displayed by fusion with coat proteins of filamentous phage. Typically, the displayed peptides are affinity-eluted against an antibody-immobilized extracellular domain of a receptor. The retained phages may be enriched by successive rounds of affinity purification and repropagation. The best binding peptides may be sequenced to identify key residues within one or more structurally related families of peptides. See, e.g., Cwirla et al. (1997), Science 276: 1696-9, in which two distinct families were identified. The peptide sequences may also suggest which residues may be safely replaced by alanine scanning or by mutagenesis at the DNA level. Mutagenesis libraries may be created and screened to further optimize the sequence of the best binders. Lowman (1997), Ann. Rev. Biophys. Biomol. Struct. 26: 401-24.

Structural analysis of protein-protein interaction may also be used to suggest peptides that mimic the binding activity of large protein ligands. In such an analysis, the crystal structure may suggest the identity and relative orientation of critical residues of the large protein ligand, from which a peptide may be designed. See, e.g., Takasaki et al. (1997), Nature Biotech. 15: 1266-70. These analytical methods may also be used to investigate the interaction between a receptor protein and peptides

selected by phage display, which may suggest further modification of the peptides to increase binding affinity.

Other methods compete with phage display in peptide research. A peptide library can be fused to the carboxyl terminus of the lac repressor and expressed in E. coli. Another E. coli-based method allows display on the cell's outer membrane by fusion with a peptidoglycan-associated lipoprotein (PAL). Hereinafter, these and related methods are collectively referred to as "E. coli display." In another method, translation of random RNA is halted prior to ribosome release, resulting in a library of polypeptides with their associated RNA still attached. Hereinafter, this and related methods are collectively referred to as "ribosome display." Other methods employ chemical linkage of peptides to RNA; see, for example, Roberts & Szostak (1997), Proc. Natl. Acad. Sci. USA, 94: 12297-303. Hereinafter, this and related methods are collectively referred to as "RNA-peptide screening." Chemically derived peptide libraries have been developed in which peptides are immobilized on stable, non-biological materials, such as polyethylene rods or solvent-permeable resins. Another chemically derived peptide library uses photolithography to scan peptides immobilized on glass slides. Hereinafter, these and related methods are collectively referred to as "chemical-peptide screening." Chemical-peptide screening may be advantageous in that it allows use of D-amino acids and other unnatural analogues, as well as non-peptide elements. Both biological and chemical methods are reviewed in Wells & Lowman (1992), Curr. Opin. Biotechnol. 3: 355-62.

Conceptually, one may discover peptide mimetics of any protein using phage display and the other methods mentioned above. These methods have been used for epitope mapping, for identification of critical amino acids in protein-protein interactions, and as leads for the discovery of new therapeutic agents. E.g., Cortese et al. (1996), Curr. Opin. Biotech. 7:

616-21. Peptide libraries are now being used most often in immunological studies, such as epitope mapping. Kreeger (1996), The Scientist 10(13): 19-20.

Of particular interest here is use of peptide libraries and other techniques in the discovery of pharmacologically active peptides. A number of such peptides identified in the art are summarized in Table 2. The peptides are described in the listed publications, each of which is hereby incorporated by reference. The pharmacologic activity of the peptides is described, and in many instances is followed by a shorthand term therefor in parentheses. Some of these peptides have been modified (e.g., to form C-terminally cross-linked dimers). Typically, peptide libraries were screened for binding to a receptor for a pharmacologically active protein (e.g., EPO receptor). In at least one instance (CTLA4), the peptide library was screened for binding to a monclonal antibody.

Table 2—Pharmacologically active peptides

Form of peptide	Binding partner/ protein of interest*	Pharmacologic activity	Reference
intrapeptide disulfide-bonded	EPO receptor	EPO-mimetic	Wrighton et al. (1996), Science 273: 458-63; U.S. Pat. No. 5,773,569, issued June 30, 1998 to Wrighton et al.
C-terminally cross-linked dimer	EPO receptor	EPO-mimetic	Livnah et al. (1996), Science 273: 464-71; Wrighton et al. (1997), Nature Biotechnology 15: 1261-5; International patent application WO 96/40772, published Dec. 19, 1996
linear	EPO receptor	EPO-mimetic	Naranda et al. (1999), Proc. Natl. Acad. Sci. USA, 96: 7569-74
linear	c-Mpl	TPO-mimetic	Cwirla et al.(1997) Science 276: 1696-9; U.S. Pat. No. 5,869,451, issued Feb. 9, 1999; U.S. Pat. No. 5,932,946, issued Aug. 3, 1999
C-terminally cross-linked dimer	c-Mpl	TPO-mimetic	Cwirla et al. (1997), Science 276: 1696-9
disulfide-linked dimer		stimulation of hematopoiesis ("G-CSF-mimetic")	Paukovits et al. (1984), Hoppe-Seylers Z. Physiol. Chem. 365: 303-11; Laerum et al. (1988), Exp. Hemat. 16: 274-80
alkylene-linked dimer		G-CSF-mimetic	Bhatnagar et al. (1996), J. Med. Chem. 39: 3814-9; Cuthbertson et al. (1997), J. Med. Chem. 40: 2876-82; King et al. (1991), Exp. Hematol. 19:481; King et al. (1995), Blood 86 (Suppl. 1): 309a
linear	IL-1 receptor	inflammatory and autoimmune diseases ("IL-1 antagonist" or "IL-1ra-mimetic")	U.S. Pat. No. 5,608,035; U.S. Pat. No. 5,786,331; U.S. Pat. No. 5,880,096; Yanofsky et al. (1996),

* The protein listed in this column may be bound by the associated peptide (e.g., EPO receptor, IL-1 receptor) or mimicked by the associated peptide. The references listed for each clarify whether the molecule is bound by or mimicked by the peptides.

			Proc. Natl. Acad. Sci. 93: 7381-6; Akeson et al. (1996), J. Biol. Chem. 271: 30517-23; Wiekzorek et al. (1997), Pol. J. Pharmacol. 49: 107-17; Yanofsky (1996), PNAs, 93:7381-7386.
linear	Facteur thymique serique (FTS)	stimulation of lymphocytes ("FTS-mimetic")	Inagaki-Ohara et al. (1996), Cellular Immunol. 171: 30-40; Yoshida (1984), Int. J. Immunopharmacol, 6:141-6.
intrapeptide disulfide bonded	CTLA4 MAb	CTLA4-mimetic	Fukumoto et al. (1998), Nature Biotech. 16: 267-70
exocyclic	TNF-α receptor	TNF-α antagonist	Takasaki et al. (1997), Nature Biotech. 15:1266-70; WO 98/53842, published December 3, 1998
linear	TNF-α receptor	TNF-α antagonist	Chirinos-Rojas ( ), J. Imm., 5621-5626.
intrapeptide disulfide bonded	C3b	inhibition of complement activation; autoimmune diseases ("C3b-antagonist")	Sahu et al. (1996), J. Immunol. 157: 884-91; Morikis et al. (1998), Protein Sci. 7: 619-27
linear	vinculin	cell adhesion processes—cell growth, differentiation, wound healing, tumor metastasis ("vinculin binding")	Adey et al. (1997), Biochem. J. 324: 523-8
linear	C4 binding protein (C4BP)	anti-thrombotic	Linse et al. (1997), J. Biol. Chem. 272: 14658-65
linear	urokinase receptor	processes associated with urokinase interaction with its receptor (e.g., angiogenesis, tumor cell invasion and metastasis); ("UKR antagonist")	Goodson et al. (1994), Proc. Natl. Acad. Sci. 91: 7129-33; International application WO 97/35969, published October 2, 1997
linear	Mdm2, Hdm2	Inhibition of inactivation of p53 mediated by Mdm2 or hdm2; anti-tumor ("Mdm/hdm antagonist")	Picksley et al. (1994), Oncogene 9: 2523-9; Bottger et al. (1997) J. Mol. Biol. 269: 744-56; Bottger et al. (1996), Oncogene 13: 2141-7
linear	$p21^{WAF1}$	anti-tumor by mimicking the activity of $p21^{WAF1}$	Ball et al. (1997), Curr. Biol. 7: 71-80
linear	farnesyl	anti-cancer by preventing	Gibbs et al. (1994), Cell

[b] FTS is a thymic hormone mimicked by the molecule of this invention rather than a receptor bound by the molecule of this invention.

	transferase	activation of ras oncogene	77:175-178
linear	Ras effector domain	anti-cancer by inhibiting biological function of the ras oncogene	Moodie et al. (1994), Trends Genet 10: 44-48 Rodriguez et al. (1994), Nature 370:527-532
linear	SH2/SH3 domains	anti-cancer by inhibiting tumor growth with activated tyrosine kinases	Pawson et al (1993), Curr. Biol. 3:434-432 Yu et al. (1994), Cell 76:933-945
linear	$p16^{INK4}$	anti-cancer by mimicking activity of p16; e.g., inhibiting cyclin D-Cdk complex ("p16-mimetic")	Fåhraeus et al. (1996), Curr. Biol. 6:84-91
linear	Src, Lyn	inhibition of Mast cell activation, IgE-related conditions, type I hypersensitivity ("Mast cell antagonist")	Stauffer et al. (1997), Biochem. 36: 9388-94
linear	Mast cell protease	treatment of inflammatory disorders mediated by release of tryptase-6 ("Mast cell protease inhibitors")	International application WO 98/33812, published August 6, 1998
linear	SH3 domains	treatment of SH3-mediated disease states ("SH3 antagonist")	Rickles et al. (1994), EMBO J. 13: 5598-5604; Sparks et al. (1994), J. Biol. Chem. 269: 23853-6; Sparks et al. (1996), Proc. Natl. Acad. Sci. 93: 1540-4
linear	HBV core antigen (HBcAg)	treatment of HBV viral infections ("anti-HBV")	Dyson & Muray (1995), Proc. Natl. Acad. Sci. 92: 2194-8
linear	selectins	neutrophil adhesion; inflammatory diseases ("selectin antagonist")	Martens et al. (1995), J. Biol. Chem. 270: 21129-36; European patent application EP 0 714 912, published June 5, 1996
linear, cyclized	calmodulin	calmodulin antagonist	Pierce et al. (1995), Molec. Diversity 1: 259-65; Dedman et al. (1993), J. Biol. Chem. 268: 23025-30; Adey & Kay (1996), Gene 169: 133-4
linear, cyclized-	integrins	tumor-homing; treatment for conditions related to integrin-mediated cellular events, including platelet aggregation, thrombosis, wound healing, osteoporosis, tissue repair, angiogenesis (e.g.,	International applications WO 95/14714, published June 1, 1995; WO 97/08203, published March 6, 1997; WO 98/10795, published March 19, 1998; WO 99/24462, published May

		for treatment of cancer), and tumor invasion ("integrin-binding")	20, 1999; Kraft et al. (1999), J. Biol. Chem. 274: 1979-1985
cyclic, linear	fibronectin and extracellular matrix components of T cells and macrophages	treatment of inflammatory and autoimmune conditions	WO 98/09985, published March 12, 1998
linear	somatostatin and cortistatin	treatment or prevention of hormone-producing tumors, acromegaly, giantism, dementia, gastric ulcer, tumor growth, inhibition of hormone secretion, modulation of sleep or neural activity	European patent application 0 911 393, published April 28, 1999
linear	bacterial lipopolysac-charide	antibiotic; septic shock; disorders modulatable by CAP37	U.S. Pat. No. 5,877,151, issued March 2, 1999
linear or cyclic, including D-amino acids	pardaxin, mellitin	antipathogenic	WO 97/31019, published 28 August 1997
linear, cyclic	VIP	impotence, neurodegenerative disorders	WO 97/40070, published October 30, 1997
linear	CTLs	cancer	EP 0 770 624, published May 2, 1997
linear	THF-gamma2		Burnstein (1988), Biochem., 27:4066-71.
linear	Amylin		Cooper (1987), Proc. Natl. Acad. Sci., 84:8628-32.
linear	Adrenomedullin		Kitamura (1993), BBRC, 192:553-60.
cyclic, linear	VEGF	anti-angiogenic; cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis ("VEGF antagonist")	Fairbrother (1998), Biochem., 37:17754-17764.
cyclic	MMP	inflammation and autoimmune disorders; tumor growth ("MMP inhibitor")	Koivunen (1999), Nature Biotech., 17:768-774.
	HGH fragment		U.S. Pat. No. 5,869,452
	Echistatin	inhibition of platelet aggregation	Gan (1988), J. Biol. Chem., 263:19827-32.
linear	SLE autoantibody	SLE	WO 96/30057, published October 3, 1996
	GD1alpha	suppression of tumor metastasis	Ishikawa et al. (1998), FEBS Lett. 441 (1): 20-4
	antiphospholipid	endothelial cell activation ,	Blank et al. (1999), Proc.

	beta-2-glycoprotein-I (β2GPI) antibodies	antiphospholipid syndrome (APS), thromboembolic phenomena, thrombocytopenia, and recurrent fetal loss	Natl. Acad. Sci. USA 96: 5164-8
linear	T Cell Receptor beta chain	diabetes	WO 96/11214, published April 18, 1996

Peptides identified by peptide library screening have been regarded as "leads" in development of therapeutic agents rather than as therapeutic agents themselves. Like other proteins and peptides, they would be rapidly removed in vivo either by renal filtration, cellular clearance mechanisms in the reticuloendothelial system, or proteolytic degradation. Francis (1992), Focus on Growth Factors 3: 4-11. As a result, the art presently uses the identified peptides to validate drug targets or as scaffolds for design of organic compounds that might not have been as easily or as quickly identified through chemical library screening. Lowman (1997), Ann. Rev. Biophys. Biomol. Struct. 26: 401-24; Kay et al. (1998), Drug Disc. Today 3: 370-8. The art would benefit from a process by which such peptides could more readily yield therapeutic agents.

Summary of the Invention

The present invention concerns a process by which the in vivo half-life of one or more biologically active peptides is increased by fusion with a vehicle. In this invention, pharmacologically active compounds are prepared by a process comprising:

a) selecting at least one peptide that modulates the activity of a protein of interest; and

b) preparing a pharmacologic agent comprising at least one vehicle covalently linked to at least one amino acid sequence of the selected peptide.

The preferred vehicle is an Fc domain. The peptides screened in step (a) are preferably expressed in a phage display library. The vehicle and the

peptide may be linked through the N- or C-terminus of the peptide or the vehicle, as described further below. Derivatives of the above compounds (described below) are also encompassed by this invention.

The compounds of this invention may be prepared by standard synthetic methods, recombinant DNA techniques, or any other methods of preparing peptides and fusion proteins. Compounds of this invention that encompass non-peptide portions may be synthesized by standard organic chemistry reactions, in addition to standard peptide chemistry reactions when applicable.

The primary use contemplated is as therapeutic or prophylactic agents. The vehicle-linked peptide may have activity comparable to—or even greater than—the natural ligand mimicked by the peptide. In addition, certain natural ligand-based therapeutic agents might induce antibodies against the patient's own endogenous ligand; the vehicle-linked peptide avoids this pitfall by having little or typically no sequence identity with the natural ligand.

Although mostly contemplated as therapeutic agents, compounds of this invention may also be useful in screening for such agents. For example, one could use an Fc-peptide (e.g., Fc-SH2 domain peptide) in an assay employing anti-Fc coated plates. The vehicle, especially Fc, may make insoluble peptides soluble and thus useful in a number of assays.

The compounds of this invention may be used for therapeutic or prophylactic purposes by formulating them with appropriate pharmaceutical carrier materials and administering an effective amount to a patient, such as a human (or other mammal) in need thereof. Other related aspects are also included in the instant invention.

Numerous additional aspects and advantages of the present invention will become apparent upon consideration of the figures and detailed description of the invention.

Brief Description of the Figures

Figure 1 shows a schematic representation of an exemplary process of the invention. In this preferred process, the vehicle is an Fc domain, which is linked to the peptide covalently by expression from a DNA construct encoding both the Fc domain and the peptide. As noted in Figure 1, the Fc domains spontaneously form a dimer in this process.

Figure 2 shows exemplary Fc dimers that may be derived from an IgG1 antibody. "Fc" in the figure represents any of the Fc variants within the meaning of "Fc domain" herein. "X^1" and "X^2" represent peptides or linker-peptide combinations as defined hereinafter. The specific dimers are as follows:

A, D: Single disulfide-bonded dimers. IgG1 antibodies typically have two disulfide bonds at the hinge region between the constant and variable domains. The Fc domain in Figures 2A and 2 D may be formed by truncation between the two disulfide bond sites or by substitution of a cysteinyl residue with an unreactive residue (e.g., alanyl). In Figure 2A, the Fc domain is linked at the amino terminus of the peptides; in 2D, at the carboxyl terminus.

B, E: Doubly disulfide-bonded dimers. This Fc domain may be formed by truncation of the parent antibody to retain both cysteinyl residues in the Fc domain chains or by expression from a construct including a sequence encoding such an Fc domain. In Figure 2B, the Fc domain is linked at the amino terminus of the peptides; in 2E, at the carboxyl terminus.

C, F: Noncovalent dimers. This Fc domain may be formed by elimination of the cysteinyl residues by either truncation or substitution. One may desire to eliminate the cysteinyl residues to avoid impurities formed by reaction of the cysteinyl residue with cysteinyl residues of other

proteins present in the host cell. The noncovalent bonding of the Fc domains is sufficient to hold together the dimer.
Other dimers may be formed by using Fc domains derived from different types of antibodies (e.g., IgG2, IgM).

Figure 3 shows the structure of preferred compounds of the invention that feature tandem repeats of the pharmacologically active peptide. Figure 3A shows a single chain molecule and may also represent the DNA construct for the molecule. Figure 3B shows a dimer in which the linker-peptide portion is present on only one chain of the dimer. Figure 3C shows a dimer having the peptide portion on both chains. The dimer of Figure 3C will form spontaneously in certain host cells upon expression of a DNA construct encoding the single chain shown in Figure 3A. In other host cells, the cells could be placed in conditions favoring formation of dimers or the dimers can be formed in vitro.

Figure 4 shows exemplary nucleic acid and amino acid sequences (SEQ ID NOS: 1 and 2, respectively) of human IgG1 Fc that may be used in this invention.

Figure 5 shows a synthetic scheme for the preparation of PEGylated peptide 19 (SEQ ID NO: 3).

Figure 6 shows a synthetic scheme for the preparation of PEGylated peptide 20 (SEQ ID NO: 4).

Figure 7 shows the nucleotide and amino acid sequences (SEQ ID NOS: 5 and 6, respectively) of the molecule identified as "Fc-TMP" in Example 2 hereinafter.

Figure 8 shows the nucleotide and amino acid sequences (SEQ. ID. NOS: 7 and 8, respectively) of the molecule identified as "Fc-TMP-TMP" in Example 2 hereinafter.

Figure 9 shows the nucleotide and amino acid sequences (SEQ. ID. NOS: 9 and 10, respectively) of the molecule identified as "TMP-TMP-Fc" in Example 2 hereinafter.

Figure 10 shows the nucleotide and amino acid sequences (SEQ. ID. NOS: 11 and 12, respectively) of the molecule identified as "TMP-Fc" in Example 2 hereinafter.

Figure 11 shows the number of platelets generated in vivo in normal female BDF1 mice treated with one 100 μg/kg bolus injection of various compounds, with the terms defined as follows.

PEG-MGDF: 20 kD average molecular weight PEG attached by reductive amination to the N-terminal amino group of amino acids 1-163 of native human TPO, which is expressed in E. coli (so that it is not glycosylated);

TMP: the TPO-mimetic peptide having the amino acid sequence IEGPTLRQWLAARA (SEQ ID NO: 13);

TMP-TMP: the TPO-mimetic peptide having the amino acid sequence IEGPTLRQWLAARA-GGGGGGGG-IEGPTLRQWLAARA (SEQ ID NO: 14);

PEG-TMP-TMP: the peptide of SEQ ID NO: 14, wherein the PEG group is a 5 kD average molecular weight PEG attached as shown in Figure 6;

Fc-TMP-TMP: the compound of SEQ ID NO: 8 (Figure 8) dimerized with an identical second monomer (i.e., Cys residues 7 and 10 are bound to the corresponding Cys residues in the second monomer to form a dimer, as shown in Figure 2); and

TMP-TMP-Fc is the compound of SEQ ID NO: 10 (Figure 9) dimerized in the same way as TMP-TMP-Fc except that the Fc domain is attached at the C-terminal end rather than the N-terminal end of the TMP-TMP peptide.

Figure 12 shows the number of platelets generated in vivo in normal BDF1 mice treated with various compounds delivered via implanted osmotic pumps over a 7-day period. The compounds are as defined for Figure 7.

Figure 13 shows the nucleotide and amino acid sequences (SEQ. ID. NOS: 15 and 16, respectively) of the molecule identified as "Fc-EMP" in Example 3 hereinafter.

Figure 14 shows the nucleotide and amino acid sequences (SEQ ID NOS: 17 and 18, respectively) of the molecule identified as "EMP-Fc" in Example 3 hereinafter.

Figure 15 shows the nucleotide and amino acid sequences (SEQ ID NOS:19 and 20, respectively) of the molecule identified as "EMP-EMP-Fc" in Example 3 hereinafter.

Figure 16 shows the nucleotide and amino acid sequences (SEQ ID NOS: 21 and 22, respectively) of the molecule identified as "Fc-EMP-EMP" in Example 3 hereinafter.

Figures 17A and 17B show the DNA sequence (SEQ ID NO: 23) inserted into pCFM1656 between the unique AatII (position #4364 in pCFM1656) and SacII (position #4585 in pCFM1656) restriction sites to form expression plasmid pAMG21 (ATCC accession no. 98113).

Figure 18A shows the hemoglobin, red blood cells, and hematocrit generated in vivo in normal female BDF1 mice treated with one 100 μg/kg bolus injection of various compounds. Figure 18B shows the same results with mice treated with 100 μg/kg per day delivered ~~the same dose~~ by 7-day micro-osmotic pump with the EMPs delivered at 100 μg/kg, rhEPO at 30U/mouse. (In both experiments, neutrophils, lymphocytes, and platelets were unaffected.) In these figures, the terms are defined as follows.

Fc-EMP: the compound of SEQ ID NO: 16 (Figure 13) dimerized with an identical second monomer (i.e., Cys residues 7 and 10 are

bound to the corresponding Cys residues in the second monomer to form a dimer, as shown in Figure 2);

EMP-Fc: the compound of SEQ ID NO: 18 (Figure 14) dimerized in the same way as Fc-EMP except that the Fc domain is attached at the C-terminal end rather than the N-terminal end of the EMP peptide.

"EMP-EMP-Fc" refers to a tandem repeat of the same peptide (SEQ ID NO: 20) attached to the same Fc domain by the carboxyl terminus of the peptides. "Fc-EMP-EMP" refers to the same tandem repeat of the peptide but with the same Fc domain attached at the amino terminus of the tandem repeat. All molecules are expressed in E. coli and so are not glycosylated.

Figures 19A and 19B show the nucleotide and amino acid sequences (SEQ ID NOS: 1055 and 1056) of the Fc-TNF-α inhibitor fusion molecule described in Example 4 hereinafter.

Figures 20A and 20B show the nucleotide and amino acid sequences (SEQ ID NOS: 1057 and 1058) of the TNF-α inhibitor-Fc fusion molecule described in Example 4 hereinafter.

Figures 21A and 21B show the nucleotide and amino acid sequences (SEQ ID NOS: 1059 and 1060) of the Fc-IL-1 antagonist fusion molecule described in Example 5 hereinafter.

Figures 22A and 22B show the nucleotide and amino acid sequences (SEQ ID NOS: 1061 and 1062) of the IL-1 antagonist-Fc fusion molecule described in Example 5 hereinafter.

Figures 23A, 23B, and 23C show the nucleotide and amino acid sequences (SEQ ID NOS: 1063 and 1064) of the Fc-VEGF antagonist fusion molecule described in Example 6 hereinafter.

Figures 24A and 24B show the nucleotide and amino acid sequences (SEQ ID NOS: 1065 and 1066) of the VEGF antagonist-Fc fusion molecule described in Example 6 hereinafter.

Figures 25A and 25B show the nucleotide and amino acid sequences (SEQ ID NOS: 1067 and 1068) of the Fc-MMP inhibitor fusion molecule described in Example 7 hereinafter.

Figures 26A and 26B show the nucleotide and amino acid sequences (SEQ ID NOS: 1069 and 1070) of the MMP inhibitor-Fc fusion molecule described in Example 7 hereinafter.

Detailed Description of the Invention

Definition of Terms

The terms used throughout this specification are defined as follows, unless otherwise limited in specific instances.

The term "comprising" means that a compound may include additional amino acids on either or both of the N- or C- termini of the given sequence. Of course, these additional amino acids should not significantly interfere with the activity of the compound.

The term "vehicle" refers to a molecule that prevents degradation and/or increases half-life, reduces toxicity, reduces immunogenicity, or increases biological activity of a therapeutic protein. Exemplary vehicles include an Fc domain (which is preferred) as well as a linear polymer (e.g., polyethylene glycol (PEG), polylysine, dextran, etc.); a branched-chain polymer (see, for example, U.S. Patent No. 4,289,872 to Denkenwalter et al., issued September 15, 1981; 5,229,490 to Tam, issued July 20, 1993; WO 93/21259 by Frechet et al., published 28 October 1993); a lipid; a cholesterol group (such as a steroid); a carbohydrate or oligosaccharide; or any natural or synthetic protein, polypeptide or peptide that binds to a salvage receptor. Vehicles are further described hereinafter.

The term "native Fc" refers to molecule or sequence comprising the sequence of a non-antigen-binding fragment resulting from digestion of whole antibody, whether in monomeric or multimeric form. The original immunoglobulin source of the native Fc is preferably of human origin and may be any of the immunoglobulins, although IgG1 and IgG2 are preferred. Native Fc's are made up of monomeric polypeptides that may be linked into dimeric or multimeric forms by covalent (i.e., disulfide bonds) and non-covalent association. The number of intermolecular disulfide bonds between monomeric subunits of native Fc molecules ranges from 1 to 4 depending on class (e.g., IgG, IgA, IgE) or subclass (e.g., IgG1, IgG2, IgG3, IgA1, IgGA2). One example of a native Fc is a disulfide-bonded dimer resulting from papain digestion of an IgG (see Ellison et al. (1982), Nucleic Acids Res. 10: 4071-9). The term "native Fc" as used herein is generic to the monomeric, dimeric, and multimeric forms.

The term "Fc variant" refers to a molecule or sequence that is modified from a native Fc but still comprises a binding site for the salvage receptor, FcRn. International applications WO 97/34631 (published 25 September 1997) and WO 96/32478 describe exemplary Fc variants, as well as interaction with the salvage receptor, and are hereby incorporated by reference. Thus, the term "Fc variant" comprises a molecule or sequence that is humanized from a non-human native Fc. Furthermore, a native Fc comprises sites that may be removed because they provide structural features or biological activity that are not required for the fusion molecules of the present invention. Thus, the term "Fc variant" comprises a molecule or sequence that lacks one or more native Fc sites or residues that affect or are involved in (1) disulfide bond formation, (2) incompatibility with a selected host cell (3) N-terminal heterogeneity upon expression in a selected host cell, (4) glycosylation, (5) interaction with complement, (6) binding to an Fc receptor other than a salvage receptor, or

(7) antibody-dependent cellular cytotoxicity (ADCC). Fc variants are described in further detail hereinafter.

The term "Fc domain" encompasses native Fc and Fc variant molecules and sequences as defined above. As with Fc variants and native Fc's, the term "Fc domain" includes molecules in monomeric or multimeric form, whether digested from whole antibody or produced by other means.

The term "multimer" as applied to Fc domains or molecules comprising Fc domains refers to molecules having two or more polypeptide chains associated covalently, noncovalently, or by both covalent and non-covalent interactions. IgG molecules typically form dimers; IgM, pentamers; IgD, dimers; and IgA, monomers, dimers, trimers, or tetramers. Multimers may be formed by exploiting the sequence and resulting activity of the native Ig source of the Fc or by derivatizing (as defined below) such a native Fc.

The term "dimer" as applied to Fc domains or molecules comprising Fc domains refers to molecules having two polypeptide chains associated covalently or non-covalently. Thus, exemplary dimers within the scope of this invention are as shown in Figure 2.

The terms "derivatizing" and "derivative" or "derivatized" comprise processes and resulting compounds respectively in which (1) the compound has a cyclic portion; for example, cross-linking between cysteinyl residues within the compound; (2) the compound is cross-linked or has a cross-linking site; for example, the compound has a cysteinyl residue and thus forms cross-linked dimers in culture or <u>in vivo</u>; (3) one or more peptidyl linkage is replaced by a non-peptidyl linkage; (4) the N-terminus is replaced by $-NRR^1$, $NRC(O)R^1$, $-NRC(O)OR^1$, $-NRS(O)_2R^1$, -NHC(O)NHR, a succinimide group, or substituted or unsubstituted benzyloxycarbonyl-NH-, wherein R and R^1 and the ring substituents are

as defined hereinafter; (5) the C-terminus is replaced by $-C(O)R^2$ or $-NR^3R^4$ wherein R^2, R^3 and R^4 are as defined hereinafter; and (6) compounds in which individual amino acid moieties are modified through treatment with agents capable of reacting with selected side chains or terminal residues. Derivatives are further described hereinafter.

The term "peptide" refers to molecules of 2 to 40 amino acids, with molecules of 3 to 20 amino acids preferred and those of 6 to 15 amino acids most preferred. Exemplary peptides may be randomly generated by any of the methods cited above, carried in a peptide library (e.g., a phage display library), or derived by digestion of proteins.

The term "randomized" as used to refer to peptide sequences refers to fully random sequences (e.g., selected by phage display methods) and sequences in which one or more residues of a naturally occurring molecule is replaced by an amino acid residue not appearing in that position in the naturally occurring molecule. Exemplary methods for identifying peptide sequences include phage display, E. coli display, ribosome display, RNA-peptide screening, chemical screening, and the like.

The term "pharmacologically active" means that a substance so described is determined to have activity that affects a medical parameter (e.g., blood pressure, blood cell count, cholesterol level) or disease state (e.g., cancer, autoimmune disorders). Thus, pharmacologically active peptides comprise agonistic or mimetic and antagonistic peptides as defined below.

The terms "-mimetic peptide" and "-agonist peptide" refer to a peptide having biological activity comparable to a protein (e.g., EPO, TPO, G-CSF) that interacts with a protein of interest. These terms further include peptides that indirectly mimic the activity of a protein of interest, such as by potentiating the effects of the natural ligand of the protein of interest; see, for example, the G-CSF-mimetic peptides listed in Tables 2

and 7. Thus, the term "EPO-mimetic peptide" comprises any peptides that can be identified or derived as described in Wrighton et al. (1996), Science 273: 458-63, Naranda et al. (1999), Proc. Natl. Acad. Sci. USA 96: 7569-74, or any other reference in Table 2 identified as having EPO-mimetic subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The term "TPO-mimetic peptide" comprises peptides that can be identified or derived as described in Cwirla et al. (1997), Science 276: 1696-9 , U.S. Pat. Nos. 5,869,451 and 5,932,946 and any other reference in Table 2 identifed as having TPO-mimetic subject matter, as well as the U.S. patent application, "Thrombopoietic Compounds," filed on even date herewith and hereby incorporated by reference. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The term "G-CSF-mimetic peptide" comprises any peptides that can be identified or described in Paukovits et al. (1984), Hoppe-Seylers Z. Physiol. Chem. 365: 303-11 or any of the references in Table 2 identified as having G-CSF-mimetic subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The term "CTLA4-mimetic peptide" comprises any peptides that can be identified or derived as described in Fukumoto et al. (1998), Nature Biotech. 16: 267-70. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually

disclosed therein by following the disclosed procedures with different peptide libraries.

The term "-antagonist peptide" or "inhibitor peptide" refers to a peptide that blocks or in some way interferes with the biological activity of the associated protein of interest, or has biological activity comparable to a known antagonist or inhibitor of the associated protein of interest. Thus, the term "TNF-antagonist peptide" comprises peptides that can be identified or derived as described in Takasaki et al. (1997), Nature Biotech. 15: 1266-70 or any of the references in Table 2 identified as having TNF-antagonistic subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The terms "IL-1 antagonist" and "IL-1ra-mimetic peptide" comprises peptides that inhibit or down-regulate activation of the IL-1 receptor by IL-1. IL-1 receptor activation results from formation of a complex among IL-1, IL-1 receptor, and IL-1 receptor accessory protein. IL-1 antagonist or IL-1ra-mimetic peptides bind to IL-1, IL-1 receptor, or IL-1 receptor accessory protein and obstruct complex formation among any two or three components of the complex. Exemplary IL-1 antagonist or IL-1ra-mimetic peptides can be identified or derived as described in U.S. Pat. Nos. 5,608,035, 5,786,331, 5,880,096, or any of the references in Table 2 identified as having IL-1ra-mimetic or IL-1 antagonistic subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The term "VEGF-antagonist peptide" comprises peptides that can be identified or derived as described in Fairbrother (1998), Biochem. 37:

17754-64, and in any of the references in Table 2 identified as having VEGF-antagonistic subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

The term "MMP inhibitor peptide" comprises peptides that can be identified or derived as described in Koivunen (1999), Nature Biotech. 17: 768-74 and in any of the references in Table 2 identified as having MMP inhibitory subject matter. Those of ordinary skill in the art appreciate that each of these references enables one to select different peptides than actually disclosed therein by following the disclosed procedures with different peptide libraries.

Additionally, physiologically acceptable salts of the compounds of this invention are also encompassed herein. By "physiologically acceptable salts" is meant any salts that are known or later discovered to be pharmaceutically acceptable. Some specific examples are: acetate; trifluoroacetate; hydrohalides, such as hydrochloride and hydrobromide; sulfate; citrate; tartrate; glycolate; and oxalate.

Structure of compounds

In General. In the compositions of matter prepared in accordance with this invention, the peptide may be attached to the vehicle through the peptide's N-terminus or C-terminus. Thus, the vehicle-peptide molecules of this invention may be described by the following formula I:

I

$$(X^1)_a\text{-}F^1\text{-}(X^2)_b$$

wherein:

F^1 is a vehicle (preferably an Fc domain);

X^1 and X^2 are each independently selected from $-(L^1)_c\text{-}P^1$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3$, and $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3\text{-}(L^4)_f\text{-}P^4$

P^1, P^2, P^3, and P^4 are each independently sequences of pharmacologically active peptides;

L^1, L^2, L^3, and L^4 are each independently linkers; and

a, b, c, d, e, and f are each independently 0 or 1, provided that at least one of a and b is 1.

Thus, compound I comprises preferred compounds of the formulae

II

$$X^1\text{-}F^1$$

and multimers thereof wherein F^1 is an Fc domain and is attached at the C-terminus of X^1;

III

$$F^1\text{-}X^2$$

and multimers thereof wherein F^1 is an Fc domain and is attached at the N-terminus of X^2;

IV

$$F^1\text{-}(L^1)_c\text{-}P^1$$

and multimers thereof wherein F^1 is an Fc domain and is attached at the N-terminus of $-(L^1)_c\text{-}P^1$; and

V

$$F^1\text{-}(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2$$

and multimers thereof wherein F^1 is an Fc domain and is attached at the N-terminus of $-L^1\text{-}P^1\text{-}L^2\text{-}P^2$.

Peptides. Any number of peptides may be used in conjunction with the present invention. Of particular interest are peptides that mimic the activity of EPO, TPO, growth hormone, G-CSF, GM-CSF, IL-1ra, leptin, CTLA4, TRAIL, TGF-α, and TGF-β. Peptide antagonists are also of interest, particularly those antagonistic to the activity of TNF, leptin, any of the interleukins (IL-1, 2, 3, ...), and proteins involved in complement activation (e.g., C3b). Targeting peptides are also of interest, including

tumor-homing peptides, membrane-transporting peptides, and the like. All of these classes of peptides may be discovered by methods described in the references cited in this specification and other references.

Phage display, in particular, is useful in generating peptides for use in the present invention. It has been stated that affinity selection from libraries of random peptides can be used to identify peptide ligands for any site of any gene product. Dedman et al. (1993), J. Biol. Chem. 268: 23025-30. Phage display is particularly well suited for identifying peptides that bind to such proteins of interest as cell surface receptors or any proteins having linear epitopes. Wilson et al. (1998), Can. J. Microbiol. 44: 313-29; Kay et al. (1998), Drug Disc. Today 3: 370-8. Such proteins are extensively reviewed in Herz et al. (1997), J. Receptor & Signal Transduction Res. 17(5): 671-776, which is hereby incorporated by reference. Such proteins of interest are preferred for use in this invention.

A particularly preferred group of peptides are those that bind to cytokine receptors. Cytokines have recently been classified according to their receptor code. See Inglot (1997), Archivum Immunologiae et Therapiae Experimentalis 45: 353-7, which is hereby incorporated by reference. Among these receptors, most preferred are the CKRs (family I in Table 3). The receptor classification appears in Table 3.

Table 3—Cytokine Receptors Classified by Receptor Code

Cytokines (ligands)		Receptor Type	
family	subfamily	family	subfamily
I. Hematopoietic cytokines	1. IL-2, IL-4, IL-7, IL-9, IL-13, IL-15	I. Cytokine R (CKR)	1. shared γCr
	2. IL-3, IL-5, GM-CSF		2. shared GP 140 βR
	3. IL-6, IL-11, IL-12, LIF, OSM, CNTF, leptin (OB)		3. 3.shared RP 130
	4. G-CSF, EPO, TPO, PRL, GH		4. "single chain" R
	5. IL-17, HVS-IL-17		5. other R[c]
II. IL-10 ligands	IL-10, BCRF-1, HSV-IL-10	II. IL-10 R	
III. Interferons	1. IFN-αl, α2, α4, m, t, IFN-β[d]	III. Interferon R	1. IFNAR
	2. IFN-γ		2. IFNGR
IV. IL-1 ligands	1. IL-1α, IL-1β, IL-1Ra	IV. IL-1R	
V. TNF ligands	1. TNF-α, TNF-β (LT), FAS1, CD40 L, CD30L, CD27 L	V. NGF/TNF R[e]	
VI. Chemokines	1. α chemokines: IL-8, GRO α, β, γ, IF-10, PF-4, SDF-1	VI. Chemokine R	1. CXCR
	2. β chemokines: MIP1α, MIP1β, MCP-1,2,3,4, RANTES, eotaxin		2. CCR
	3. γ chemokines: lymphotactin		3. CR
			4. DARC[f]

[c] IL-17R belongs to the CKR family but is not assigned to any of the 4 indicated subjamilies.
[d] Other IFN type I subtypes remain unassigned. Hematopoietic cytokines, IL-10 ligands and interferons do not possess functional intrinsic protein kinases. The signaling molecules for the cytokines are JAK's, STATs and related non-receptor molecules. IL-14, IL-16 and IL-18 have been cloned but according to the receptor code they remain unassigned.
[e] TNF receptors use multiple, distinct intracellular molecules for signal transduction including "death domain" of FAS R and 55 kDa TNF-αR that participates in their cytotoxic effects. NGF/TNF R can bind both NGF and related factors as well as TNF ligands. Chemokine receptors are G protein-coupled, seven transmembrane (7TM, serpentine) domain receptors.
[f] The Duffy blood group antigen (DARC) is an erythrocyte receptor that can bind several different chemokines. It belongs to the immunoglobulin superfamily but characteristics of its signal transduction events remain unclear.

VII. Growth factors	VII. RKF
1.1 SCF, M-CSF, PDGF-AA, AB, BB, FLT-3L, VEGF, SSV-PDGF	1. TK sub-family 1.1 IgTK III R
1.2 FGFα, FGFβ	1.2 IgTK IV R
1.3 EGF, TGF-α, VV-F19 (EGF-like)	1.3 Cysteine-rich TK-I
1.4 IGF-I, IGF-II, Insulin	1.4 Cysteine rich TK-II
1.5 NGF, BDNF, NT-3, NT-4[g]	1.5 Cysteine knot TK V
2. TGF-β1,β2,β3	2. STK subfamily[h]

Exemplary peptides for this invention appear in Tables 4 through 20 below. These peptides may be prepared by methods disclosed in the art. Single letter amino acid abbreviations are used. The X in these sequences (and throughout this specification, unless specified otherwise in a particular instance) means that any of the 20 naturally occurring amino acid residues may be present. Any of these peptides may be linked in tandem (i.e., sequentially), with or without linkers, and a few tandem-linked examples are provided in the table. Linkers are listed as "Λ" and may be any of the linkers described herein. Tandem repeats and linkers are shown separated by dashes for clarity. Any peptide containing a cysteinyl residue may be cross-linked with another Cys-containing peptide, either or both of which may be linked to a vehicle. A few cross-linked examples are provided in the table. Any peptide having more than one Cys residue may form an intrapeptide disulfide bond, as well; see, for example, EPO-mimetic peptides in Table 5. A few examples of intrapeptide disulfide-bonded peptides are specified in the table. Any of these peptides may be derivatized as described herein, and a few derivatized examples are provided in the table. Derivatized peptides in

[g] The neurotrophic cytokines can associate with NGF/TNF receptors also.

the tables are exemplary rather than limiting, as the associated underivatized peptides may be employed in this invention, as well. For derivatives in which the carboxyl terminus may be capped with an amino group, the capping amino group is shown as -NH_2. For derivatives in which amino acid residues are substituted by moieties other than amino acid residues, the substitutions are denoted by σ, which signifies any of the moieties described in Bhatnagar et al. (1996), J. Med. Chem. 39: 3814-9 and Cuthbertson et al. (1997), J. Med. Chem. 40: 2876-82, which are incorporated by reference. The J substituent and the Z substituents (Z_5, Z_6, ...Z_{40}) are as defined in U.S. Pat. Nos. 5,608,035 ,5,786,331, and 5,880,096, which are incorporated by reference. For the EPO-mimetic sequences (Table 5), the substituents X_2 through X_{11} and the integer "n" are as defined in WO 96/40772, which is incorporated by reference. The substituents "Ψ," "Θ," and "+" are as defined in Sparks et al. (1996), Proc. Natl. Acad. Sci. 93: 1540-4, which is hereby incorporated by reference. X_4, X_5, X_6, and X_7 are as defined in U.S. Pat. No. 5,773,569, which is hereby incorporated by reference, except that: for integrin-binding peptides, X_1, X_2, X_3, X_4, X_5, X_6, X_7, and X_8 are as defined in International applications WO 95/14714, published June 1, 1995 and WO 97/08203, published March 6, 1997, which are also incorporated by reference; and for VIP-mimetic peptides, X_1, X_1', X_1'', X_2, X_3, X_4, X_5, X_6 and Z and the integers m and n are as defined in WO 97/40070, published October 30, 1997, which is also incorporated by reference. Xaa and Yaa below are as defined in WO 98/09985, published March 12, 1998, which is incorporated by reference. AA_1, AA_2, AB_1, AB_2, and AC are as defined in International application WO 98/53842, published December 3, 1998, which is incorporated by reference. X^1, X^2, X^3, and X^4 in Table 17 only are as defined in European application EP 0 911

[h] STKS may encompass many other TGF-β-related factors that remain unassigned. The protein kinases are intrinsic part of the intracellular domain of receptor kinase family (RKF). The enzymes participate in the signals transmission via the receptors.

393, published April 28, 1999. Residues appearing in boldface are D-amino acids. All peptides are linked through peptide bonds unless otherwise noted. Abbreviations are listed at the end of this specification. In the "SEQ ID NO." column, "NR" means that no sequence listing is required for the given sequence.

Table 4—IL-1 antagonist peptide sequences

Sequence/structure	SEQ ID NO:
$Z_{11}Z_7Z_8QZ_5YZ_6Z_9Z_{10}$	212
XXQZ$_5$YZ$_6$XX	907
Z_7XQZ$_5$YZ$_6$XX	908
$Z_7Z_8QZ_5YZ_6Z_9Z_{10}$	909
$Z_{11}Z_7Z_8QZ_5YZ_6Z_9Z_{10}$	910
$Z_{12}Z_{13}Z_{14}Z_{15}Z_{16}Z_{17}Z_{18}Z_{19}Z_{20}Z_{21}Z_{22}Z_{11}Z_7Z_8QZ_5YZ_6Z_9Z_{10}L$	917
$Z_{23}NZ_{24}Z_{25}Z_{26}Z_{27}Z_{28}Z_{29}Z_{30}Z_{40}$	979
TANVSSFEWTPYYWQPYALPL	213
SWTDYGYWQPYALPISGL	214
ETPFTWEESNAYYWQPYALPL	215
ENTYSPNWADSMYWQPYALPL	216
SVGEDHNFWTSEYWQPYALPL	217
DGYDRWRQSGERYWQPYALPL	218
FEWTPGYWQPY	219
FEWTPGYWQHY	220
FEWTPGWYQJY	221
AcFEWTPGWYQJY	222
FEWTPGWpYQJY	223
FAWTPGYWQJY	224
FEWAPGYWQJY	225
FEWVPGYWQJY	226
FEWTPGYWQJY	227
AcFEWTPGYWQJY	228
FEWTPaWYQJY	229
FEWTPSarWYQJY	230
FEWTPGYYQPY	231
FEWTPGWWQPY	232
FEWTPNYWQPY	233
FEWTPvYWQJY	234
FEWTPecGYWQJY	235
FEWTPAibYWQJY	236
FEWTSarGYWQJY	237
FEWTPGYWQPY	238
FEWTPGYWQHY	239
FEWTPGWYQJY	240

AcFEWTPGWYQJY	241
FEWTPGW-pY-QJY	242
FAWTPGYWQJY	243
FEWAPGYWQJY	244
FEWVPGYWQJY	245
FEWTPGYWQJY	246
AcFEWTPGYWQJY	247
FEWTPAWYQJY	248
FEWTPSarWYQJY	249
FEWTPGYYQPY	250
FEWTPGWWQPY	251
FEWTPNYWQPY	252
FEWTPVYWQJY	253
FEWTPecGYWQJY	254
FEWTPAibYWQJY	255
FEWTSarGYWQJY	256
FEWTPGYWQPYALPL	257
1NapEWTPGYYQJY	258
YEWTPGYYQJY	259
FEWVPGYYQJY	260
FEWTPSYYQJY	261
FEWTPNYYQJY	262
TKPR	263
RKSSK	264
RKQDK	265
NRKQDK	266
RKQDKR	267
ENRKQDKRF	268
VTKFYF	269
VTKFY	270
VTDFY	271
SHLYWQPYSVQ	671
TLVYWQPYSLQT	672
RGDYWQPYSVQS	673
VHVYWQPYSVQT	674
RLVYWQPYSVQT	675
SRVWFQPYSLQS	676
NMVYWQPYSIQT	677
SVVFWQPYSVQT	678
TFVYWQPYALPL	679
TLVYWQPYSIQR	680
RLVYWQPYSVQR	681
SPVFWQPYSIQI	682
WIEWWQPYSVQS	683
SLIYWQPYSLQM	684
TRLYWQPYSVQR	685
RCDYWQPYSVQT	686
MRVFWQPYSVQN	687
KIVYWQPYSVQT	688
RHLYWQPYSVQR	689

ALVWWQPYSEQI	690
SRVWFQPYSLQS	691
WEQPYALPLE	692
QLVWWQPYSVQR	693
DLRYWQPYSVQV	694
ELVWWQPYSLQL	695
DLVWWQPYSVQW	696
NGNYWQPYSFQV	697
ELVYWQPYSIQR	698
ELMYWQPYSVQE	699
NLLYWQPYSMQD	700
GYEWYQPYSVQR	701
SRVWYQPYSVQR	702
LSEQYQPYSVQR	703
GGGWWQPYSVQR	704
VGRWYQPYSVQR	705
VHVYWQPYSVQR	706
QARWYQPYSVQR	707
VHVYWQPYSVQT	708
RSVYWQPYSVQR	709
TRVWFQPYSVQR	710
GRIWFQPYSVQR	711
GRVWFQPYSVQR	712
ARTWYQPYSVQR	713
ARVWWQPYSVQM	714
RLMFYQPYSVQR	715
ESMWYQPYSVQR	716
HFGWWQPYSVHM	717
ARFWWQPYSVQR	718
RLVYWQ PYAPIY	719
RLVYWQ PYSYQT	720
RLVYWQ PYSLPI	721
RLVYWQ PYSVQA	722
SRVWYQ PYAKGL	723
SRVWYQ PYAQGL	724
SRVWYQ PYAMPL	725
SRVWYQ PYSVQA	726
SRVWYQ PYSLGL	727
SRVWYQ PYAREL	728
SRVWYQ PYSRQP	729
SRVWYQ PYFVQP	730
EYEWYQ PYALPL	731
IPEYWQ PYALPL	732
SRIWWQ PYALPL	733
DPLFWQ PYALPL	734
SRQWVQ PYALPL	735
IRSWWQ PYALPL	736
RGYWQ PYALPL	737
RLLWVQ PYALPL	738
EYRWFQ PYALPL	739

DAYWVQ PYALPL	740
WSGYFQ PYALPL	741
NIEFWQ PYALPL	742
TRDWVQ PYALPL	743
DSSWYQ PYALPL	744
IGNWYQ PYALPL	745
NLRWDQ PYALPL	746
LPEFWQ PYALPL	747
DSYWWQ PYALPL	748
RSQYYQ PYALPL	749
ARFWLQ PYALPL	750
NSYFWQ PYALPL	751
RFMYWQPYSVQR	752
AHLFWQPYSVQR	753
WWQPYALPL	754
YYQPYALPL	755
YFQPYALGL	756
YWYQPYALPL	757
RWWQPYATPL	758
GWYQPYALGF	759
YWYQPYALGL	760
IWYQPYAMPL	761
SNMQPYQRLS	762
TFVYWQPY AVGLPAAETACN	763
TFVYWQPY SVQMTITGKVTM	764
TFVYWQPY SSHXXVPXGFPL	765
TFVYWQPY YGNPQWAIHVRH	766
TFVYWQPY VLLELPEGAVRA	767
TFVYWQPY VDYVWPIPIAQV	768
GWYQPYVDGWR	769
RWEQPYVKDGWS	770
EWYQPYALGWAR	771
GWWQPYARGL	772
LFEQPYAKALGL	773
GWEQPYARGLAG	774
AWVQPYATPLDE	775
MWYQPYSSQPAE	776
GWTQPYSQQGEV	777
DWFQPYSIQSDE	778
PWIQPYARGFG	779
RPLYWQPYSVQV	780
TLIYWQPYSVQI	781
RFDYWQPYSDQT	782
WHQFVQPYALPL	783
EWDS VYWQPYSVQ TLLR	784
WEQN VYWQPYSVQ SFAD	785
SDV VYWQPYSVQ SLEM	786
YYDG VYWQPYSVQ VMPA	787
SDIWYQ PYALPL	788
QRIWWQ PYALPL	789

SRIWWQ PYALPL	790
RSLYWQ PYALPL	791
TIIWEQ PYALPL	792
WETWYQ PYALPL	793
SYDWEQ PYALPL	794
SRIWCQ PYALPL	795
EIMFWQ PYALPL	796
DYVWQQ PYALPL	797
MDLLVQ WYQPYALPL	798
GSKVIL WYQPYALPL	799
RQGANI WYQPYALPL	800
GGGDEP WYQPYALPL	801
SQLERT WYQPYALPL	802
ETWVRE WYQPYALPL	803
KKGSTQ WYQPYALPL	804
LQARMN WYQPYALPL	805
EPRSQK WYQPYALPL	806
VKQKWR WYQPYALPL	807
LRRHDV WYQPYALPL	808
RSTASI WYQPYALPL	809
ESKEDQ WYQPYALPL	810
EGLTMK WYQPYALPL	811
EGSREG WYQPYALPL	812
VIEWWQ PYALPL	813
VWYWEQ PYALPL	814
ASEWWQ PYALPL	815
FYEWWQ PYALPL	816
EGWWVQ PYALPL	817
WGEWLQ PYALPL	818
DYVWEQ PYALPL	819
AHTWWQ PYALPL	820
FIEWFQ PYALPL	821
WLAWEQ PYALPL	822
VMEWWQ PYALPL	823
ERMWQ PYALPL	824
NXXWXX PYALPL	825
WGNWYQ PYALPL	826
TLYWEQ PYALPL	827
VWRWEQ PYALPL	828
LLWTQ PYALPL	829
SRIWXX PYALPL	830
SDIWYQ PYALPL	831
WGYYXX PYALPL	832
TSGWYQ PYALPL	833
VHPYXX PYALPL	834
EHSYFQ PYALPL	835
XXIWYQ PYALPL	836
AQLHSQ PYALPL	837
WANWFQ PYALPL	838
SRLYSQ PYALPL	839

GVTFSQ PYALPL	840
SIVWSQ PYALPL	841
SRDLVQ PYALPL	842
HWGH VYWQPYSVQ DDLG	843
SWHS VYWQPYSVQ SVPE	844
WRDS VYWQPYSVQ PESA	845
TWDA VYWQPYSVQ KWLD	846
TPPW VYWQPYSVQ SLDP	847
YWSS VYWQPYSVQ SVHS	848
YWY QPY ALGL	849
YWY QPY ALPL	850
EWI QPY ATGL	851
NWE QPY AKPL	852
AFY QPY ALPL	853
FLY QPY ALPL	854
VCK QPY LEWC	855
ETPFTWEESNAYYWQPYALPL	856
QGWLTWQDSVDMYWQPYALPL	857
FSEAGYTWPENTYWQPYALPL	858
TESPGGLDWAKIYWQPYALPL	859
DGYDRWRQSGERYWQPYALPL	860
TANVSSFEWTPGYWQPYALPL	861
SVGEDHNFWTSE YWQPYALPL	862
MNDQTSEVSTFP YWQPYALPL	863
SWSEAFEQPRNL YWQPYALPL	864
QYAEPSALNDWG YWQPYALPL	865
NGDWATADWSNY YWQPYALPL	866
THDEHI YWQPYALPL	867
MLEKTYTTWTPG YWQPYALPL	868
WSDPLTRDADL YWQPYALPL	869
SDAFTTQDSQAM YWQPYALPL	870
GDDAAWRTDSLT YWQPYALPL	871
AIIRQLYRWSEM YWQPYALPL	872
ENTYSPNWADSM YWQPYALPL	873
MNDQTSEVSTFP YWQPYALPL	874
SVGEDHNFWTSE YWQPYALPL	875
QTPFTWEESNAY YWQPYALPL	876
ENPFTWQESNAY YWQPYALPL	877
VTPFTWEDSNVF YWQPYALPL	878
QIPFTWEQSNAY YWQPYALPL	879
QAPLTWQESAAY YWQPYALPL	880
EPTFTWEESKAT YWQPYALPL	881
TTTLTWEESNAY YWQPYALPL	882
ESPLTWEESSAL YWQPYALPL	883
ETPLTWEESNAY YWQPYALPL	884
EATFTWAESNAY YWQPYALPL	885
EALFTWKESTAY YWQPYALPL	886
STP-TWEESNAY YWQPYALPL	887
ETPFTWEESNAY YWQPYALPL	888
KAPFTWEESQAY YWQPYALPL	889

STSFTWEESNAY YWQPYALPL	890
DSTFTWEESNAY YWQPYALPL	891
YIPFTWEESNAY YWQPYALPL	892
QTAFTWEESNAY YWQPYALPL	893
ETLFTWEESNAT YWQPYALPL	894
VSSFTWEESNAY YWQPYALPL	895
QPYALPL	896
Py-1-NapPYQJYALPL	897
TANVSSFEWTPG YWQPYALPL	898
FEWTPGYWQPYALPL	899
FEWTPGYWQJYALPL	900
FEWTPGYYQJYALPL	901
ETPFTWEESNAYYWQPYALPL	902
FTWEESNAYYWQJYALPL	903
ADVL YWQPYA PVTLWV	904
GDVAE YWQPYA LPLTSL	905
SWTDYG YWQPYA LPISGL	906
FEWTPGYWQPYALPL	911
FEWTPGYWQJYALPL	912
FEWTPGWYQPYALPL	913
FEWTPGWYQJYALPL	914
FEWTPGYYQPYALPL	915
FEWTPGYYQJYALPL	916
TANVSSFEWTPGYWQPYALPL	918
SWTDYGYWQPYALPISGL	919
ETPFTWEESNAYYWQPYALPL	920
ENTYSPNWADSMYWQPYALPL	921
SVGEDHNFWTSEYWQPYALPL	922
DGYDRWRQSGERYWQPYALPL	923
FEWTPGYWQPYALPL	924
FEWTPGYWQPY	925
FEWTPGYWQJY	926
EWTPGYWQPY	927
FEWTPGWYQJY	928
AEWTPGYWQJY	929
FAWTPGYWQJY	930
FEATPGYWQJY	931
FEWAPGYWQJY	932
FEWTAGYWQJY	933
FEWTPAYWQJY	934
FEWTPGAWQJY	935
FEWTPGYAQJY	936
FEWTPGYWQJA	937
FEWTGGYWQJY	938
FEWTPGYWQJY	939
FEWTJGYWQJY	940
FEWTPecGYWQJY	941
FEWTPAibYWQJY	942
FEWTPSarWYQJY	943
FEWTSarGYWQJY	944

FEWTPNYWQJY	945
FEWTPVYWQJY	946
FEWTVPYWQJY	947
AcFEWTPGWYQJY	948
AcFEWTPGYWQJY	949
lNap-EWTPGYYQJY	950
YEWTPGYYQJY	951
FEWVPGYYQJY	952
FEWTPGYYQJY	953
FEWTPsYYQJY	954
FEWTPnYYQJY	955
SHLY-Nap-QPYSVQM	956
TLVY-Nap-QPYSLQT	957
RGDY-Nap-QPYSVQS	958
NMVY-Nap-QPYSIQT	959
VYWQPYSVQ	960
VY-Nap-QPYSVQ	961
TFVYWQJYALPL	962
FEWTPGYYQJ-Bpa	963
XaaFEWTPGYYQJ-Bpa	964
FEWTPGY-Bpa-QJY	965
AcFEWTPGY-Bpa-QJY	966
FEWTPG-Bpa-YQJY	967
AcFEWTPG-Bpa-YQJY	968
AcFE-Bpa-TPGYYQJY	969
AcFE-Bpa-TPGYYQJY	970
Bpa-EWTPGYYQJY	971
AcBpa-EWTPGYYQJY	972
VYWQPYSVQ	973
RLVYWQPYSVQR	974
RLVY-Nap-QPYSVQR	975
RLDYWQPYSVQR	976
RLVWFQPYSVQR	977
RLVYWQPYSIQR	978
DNSSWYDSFLL	980
DNTAWYESFLA	981
DNTAWYENFLL	982
PARE DNTAWYDSFLI WC	983
TSEY DNTTWYEKFLA SQ	984
SQIP DNTAWYQSFLL HG	985
SPFI DNTAWYENFLL TY	986
EQIY DNTAWYDHFLL SY	987
TPFI DNTAWYENFLL TY	988
TYTY DNTAWYERFLM SY	989
TMTQ DNTAWYENFLL SY	990
TI DNTAWYANLVQ TYPQ	991
TI DNTAWYERFLA QYPD	992
HI DNTAWYENFLL TYTP	993
SQ DNTAWYENFLL SYKA	994
QI DNTAWYERFLL QYNA	995

NQ DNTAWYESFLL QYNT	996
TI DNTAWYENFLL NHNL	997
HY DNTAWYERFLQ QGWH	998
ETPFTWEESNAYYWQPYALPL	999
YIPFTWEESNAYYWQPYALPL	1000
DGYDRWRQSGERYWQPYALPL	1001
pY-INap-pY-QJYALPL	1002
TANVSSFEWTPGYWQPYALPL	1003
FEWTPGYWQJYALPL	1004
FEWTPGYWQPYALPLSD	1005
FEWTPGYYQJYALPL	1006
FEWTPGYWQJY	1007
AcFEWTPGYWQJY	1008
AcFEWTPGWYQJY	1009
AcFEWTPGYYQJY	1010
AcFEWTPaYWQJY	1011
AcFEWTPaWYQJY	1012
AcFEWTPaYYQJY	1013
FEWTPGYYQJYALPL	1014
FEWTPGYWQJYALPL	1015
FEWTPGWYQJYALPL	1016
TANVSSFEWTPGYWQPYALPL	1017
AcFEWTPGYWQJY	1018
AcFEWTPGWYQJY	1019
AcFEWTPGYYQJY	1020
AcFEWTPAYWQJY	1021
AcFEWTPAWYQJY	1022
AcFEWTPAYYQJY	1023

Table 5—EPO-mimetic peptide sequences

Sequence/structure	SEQ ID NO:
YXCXXGPXTWXCXP	83
YXCXXGPXTWXCXP-YXCXXGPXTWXCXP	84
YXCXXGPXTWXCXP-Λ-YXCXXGPXTWXCXP	85
YXCXXGPXTWXCXP-Λ-(ε-amine) K βA YXCXXGPXTWXCXP-Λ-(α-amine)	86 86
GGTYSCHFGPLTWVCKPQGG	87
GGDYHCRMGPLTWVCKPLGG	88
GGVYACRMGPITWVCSPLGG	89
VGNYMCHFGPITWVCRPGGG	90
GGLYLCRFGPVTWDCGYKGG	91
GGTYSCHFGPLTWVCKPQGG- GGTYSCHFGPLTWVCKPQGG	92
GGTYSCHFGPLTWVCKPQGG -Λ- GGTYSCHFGPLTWVCKPQGG	93
GGTYSCHFGPLTWVCKPQGGSSK	94
GGTYSCHFGPLTWVCKPQGGSSK- GGTYSCHFGPLTWVCKPQGGSSK	95
GGTYSCHFGPLTWVCKPQGGSSK-Λ- GGTYSCHFGPLTWVCKPQGGSSK	96
GGTYSCHFGPLTWVCKPQGGSS(ε-amine) K βA GGTYSCHFGPLTWVCKPQGGSS(α-amine)	97 97
GGTYSCHFGPLTWVCKPQGGSSK(–Λ-biotin)	98
$CX_4X_5GPX_8TWX_7C$	421
GGTYSCHGPLTWVCKPQGG	422
VGNYMAHMGPITWVCRPGG	423
GGPHHVYACRMGPLTWIC	424
GGTYSCHFGPLTWVCKPQ	425
GGLYACHMGPMTWVCQPLRG	426
TIAQYICYMGPETWECRPSPKA	427
YSCHFGPLTWVCK	428
YCHFGPLTWVC	429
$X_3X_4X_5GPX_8TWX_7X_8$	124
$YX_2X_3X_4X_5GPX_8TWX_7X_8$	461

$X_1YX_2X_3X_4X_5GPX_6TWX_7X_8X_9X_{10}X_{11}$	419
$X_1YX_2CX_4X_5GPX_6TWX_7CX_9X_{10}X_{11}$	420
GGLYLCRFGPVTWDCGYKGG	1024
GGTYSCHFGPLTWVCKPQGG	1025
GGDYHCRMGPLTWVCKPLGG	1026
VGNYMCHFGPITWVCRPGGG	1029
GGVYACRMGPITWVCSPLGG	1030
VGNYMAHMGPITWVCRPGG	1035
GGTYSCHFGPLTWVCKPQ	1036
GGLYACHMGPMTWVCQPLRG	1037
TIAQYICYMGPETWECRPSPKA	1038
YSCHFGPLTWVCK	1039
YCHFGPLTWVC	1040
SCHFGPLTWVCK	1041
$(AX_2)_nX_3X_4X_5GPX_6TWX_7X_8$	1042

Table 6—TPO-mimetic peptide sequences

Sequence/structure	SEQ ID NO:
IEGPTLRQWLAARA	13
IEGPTLRQWLAAKA	24
IEGPTLREWLAARA	25
IEGPTLRQWLAARA-Λ-IEGPTLRQWLAARA	26
IEGPTLRQWLAAKA-Λ-IEGPTLRQWLAAKA	27
IEGPTLRQCLAARA-Λ-IEGPTLRQCLAARA	28
IEGPTLRQWLAARA-Λ-K(BrAc)-Λ-IEGPTLRQWLAARA	29
IEGPTLRQWLAARA-Λ-K(PEG)-Λ-IEGPTLRQWLAARA	30
IEGPTLRQCLAARA-Λ-IEGPTLRQWLAARA | IEGPTLRQCLAARA-Λ-IEGPTLRQWLAARA	31 31
IEGPTLRQWLAARA-Λ-IEGPTLRQCLAARA | IEGPTLRQWLAARA-Λ-IEGPTLRQCLAARA	32 32
VRDQIXXXL	33
TLREWL	34
GRVRDQVAGW	35
GRVKDQIAQL	36
GVRDQVSWAL	37
ESVREQVMKY	38
SVRSQISASL	39
GVRETVYRHM	40
GVREVIVMHML	41
GRVRDQIWAAL	42
AGVRDQILIWL	43
GRVRDQIMLSL	44
GRVRDQI$(X)_1$L	45
CTLRQWLQGC	46
CTLQEFLEGC	47
CTRTEWLHGC	48
CTLREWLHGGFC	49
CTLREWVFAGLC	50
CTLRQWLILLGMC	51
CTLAEFLASGVEQC	52
CSLQEFLSHGGYVC	53
CTLREFLDPTTAVC	54
CTLKEWLVSHEVWC	55
CTLREWL$(X)_{2-5}$C	56-60
REGPTLRQWM	61
EGPTLRQWLA	62
ERGPFWAKAC	63
REGPRCVMWM	64
CGTEGPTLSTWLDC	65

CEQDGPTLLEWLKC	66
CELVGPSLMSWLTC	67
CLTGPFVTQWLYEC	68
CRAGPTLLEWLTLC	69
CADGPTLREWISFC	70
$C(X)_{1-2}EGPTLREWL(X)_{1-2}C$	71-74
GGCTLREWLHGGFCGG	75
GGCADGPTLREWISFCGG	76
GNADGPTLRQWLEGRRPKN	77
LAIEGPTLRQWLHGNGRDT	78
HGRVGPTLREWKTQVATKK	79
TIKGPTLRQWLKSREHTS	80
ISDGPTLKEWLSVTRGAS	81
SIEGPTLREWLTSRTPHS	82

Table 7—G-CSF-mimetic peptide sequences

Sequence/structure	SEQ ID NO:
EEDCK	99
EEDCK | EEDCK	99 99
EEDσK	100
EEDσK | EEDσK	100 100
pGluEDσK	101
pGluEDσK | pGluEDσK	101 101
PicSDσK	102
PicSDσK | PicSDσK	102 102
EEDCK-Λ-EEDCK	103
EEDXK-Λ–EEDXK	104

Table 8—TNF-antagonist peptide sequences

Sequence/structure	SEQ ID NO:
YCFTASENHCY	106
YCFTNSENHCY	107
YCFTRSENHCY	108
FCASENHCY	109
YCASENHCY	110
FCNSENHCY	111
FCNSENRCY	112
FCNSVENRCY	113
YCSQSVSNDCF	114
FCVSNDRCY	115
YCRKELGQVCY	116
YCKEPGQCY	117
YCRKEMGCY	118
FCRKEMGCY	119
YCWSQNLCY	120
YCELSQYLCY	121
YCWSQNYCY	122
YCWSQYLCY	123
DFLPHYKNTSLGHRP	1085
AA_1-AB_1 \ AC / AA_2-AB_2	NR

Table 9—Integrin-binding peptide sequences

Sequence/structure	SEQ ID NO:
$RX_1ETX_2WX_3$	441
$RX_1ETX_2WX_3$	442
RGDGX	443
CRGDGXC	444
$CX_1X_2RLDX_3X_4C$	445
CARRLDAPC	446
CPSRLDSPC	447
$X_1X_2X_3RGDX_4X_5X_6$	448
$CX_2CRGDCX_5C$	449
CDCRGDCFC	450
CDCRGDCLC	451
CLCRGDCIC	452
$X_1X_2DDX_4X_5X_7X_8$	453
$X_1X_2X_3DDX_4X_5X_6X_7X_8$	454
CWDDGWLC	455
CWDDLWWLC	456
CWDDGLMC	457
CWDDGWMC	458
CSWDDGWLC	459
CPDDLWWLC	460
NGR	NR
GSL	NR
RGD	NR
CGRECPRLCQSSC	1071
CNGRCVSGCAGRC	1072
CLSGSLSC	1073
RGD	NR
NGR	NR
GSL	NR
NGRAHA	1074
CNGRC	1075
CDCRGDCFC	1076
CGSLVRC	1077
DLXXL	1043
RTDLDSLRTYTL	1044
RTDLDSLRTY	1053
RTDLDSLRT	1054
RTDLDSLR	1078
GDLDLLKLRLTL	1079
GDLHSLRQLLSR	1080
RDDLHMLRLQLW	1081
SSDLHALKKRYG	1082
RGDLKQLSELTW	1083
RGDLAALSAPPV	1084

Table 10—Selectin antagonist peptide sequences

Sequence/structure	SEQ ID NO:
DITWDQLWDLMK	147
DITWDELWKIMN	148
DYTWFELWDMMQ	149
QITWAQLWNMMK	150
DMTWHDLWTLMS	151
DYSWHDLWEMMS	152
EITWDQLWEVMN	153
HVSWEQLWDIMN	154
HITWDQLWRIMT	155
RNMSWLELWEHMK	156
AEWTWDQLWHVMNPAESQ	157
HRAEWLALWEQMSP	158
KKEDWLALWRIMSV	159
ITWDQLWDLMK	160
DITWDQLWDLMK	161
DITWDQLWDLMK	162
DITWDQLWDLMK	163
CQNRYTDLVAIQNKNE	462
AENWADNEPNNKRNNED	463
RKNNKTWTWVGTKKALTNE	464
KKALTNEAENWAD	465
CQXRYTDLVAIQNKXE	466
RKXNXXWTWVGTXKXLTEE	467
AENWADGEPNNKXNXED	468
CXXXYTXLVAIQNKXE	469
RKXXXXWXWVGTXKXLTXE	470
AXNWXXXEPNNXXXED	471
XKXKTXEAXNWXX	472

Table 11—Antipathogenic peptide sequences

Sequence/structure	SEQ ID NO:
GFFALIPKIISSPLFKTLLSAVGSALSSSGGQQ	503
GFFALIPKIISSPLFKTLLSAVGSALSSSGGQE	504
GFFALIPKIISSPLFKTLLSAV	505
GFFALIPKIISSPLFKTLLSAV	506
KGFFALIPKIISSPLFKTLLSAV	507
KKGFFALIPKIISSPLFKTLLSAV	508
KKGFFALIPKIISSPLFKTLLSAV	509
GFFALIPKIIS	510
GIGAVLKVLTTGLPALISWIKRKRQQ	511
GIGAVLKVLTTGLPALISWIKRKRQQ	512
GIGAVLKVLTTGLPALISWIKRKRQQ	513
GIGAVLKVLTTGLPALISWIKR	514
AVLKVLTTGLPALISWIKR	515
KLLLLLKLLLLK	516
KLLLKLLLKLLK	517
KLLLKLKLKLLK	518
KKLLKLKLKLKK	519
KLLLKLLLKLLK	520
KLLLKLKLKLLK	521
KLLLLK	522
KLLLKLLK	523
KLLLKLKLKLLK	524
KLLLKLKLKLLK	525
KLLLKLKLKLLK	526
KAAAKAAAKAAK	527
KVVVKVVVKVVK	528
KVVVKVKVKVVK	529
KVVVKVKVKVK	530
KVVVKVKVKVVK	531
KLILKL	532
KVLHLL	533
LKLRLL	534
KPLHLL	535
KLILKLVR	536
KVFHLLHL	537
HKFRILKL	538
KPFHILHL	539
KIIIKIKIKIIK	540
KIIIKIKIKIIK	541
KIIIKIKIKIIK	542
KIPIKIKIKIPK	543
KIPIKIKIKIVK	544
RIIIRIRIRIIR	545
RIIIRIRIRIIR	546
RIIIRIRIRIIR	547
RIVIRIRIRLIR	548

RIIVRIRLRIIR	549
RIGIRLRVRIIR	550
KIVIRIRIRLIR	551
RIAVKWRLRFIK	552
KIGWKLRVRIIR	553
KKIGWLIIRVRR	554
RIVIRIRIRLIRIR	555
RIIVRIRLRIIRVR	556
RIGIRLRVRIIRRV	557
KIVIRIRARLIRIRIR	558
RIIVKIRLRIIKKIRL	559
KIGIKARVRIIRVKII	560
RIIVHIRLRIIHHIRL	561
HIGIKAHVRIIRVHII	562
RIYVKIHLRYIKKIRL	563
KIGHKARVHIIRYKII	564
RIYVKPHPRYIKKIRL	565
KPGHKARPHIIRYKII	566
KIVIRIRIRLIRIRIRKIV	567
RIIVKIRLRIIKKIRLIKK	568
KIGWKLRVRIIRVKIGRLR	569
KIVIRIRIRLIRIRIRKIVKVKRIR	570
RFAVKIRLRIIKKIRLIKKIRKRVIK	571
KAGWKLRVRIIRVKIGRLRKIGWKKRVRIK	572
RIYVKPHPRYIKKIRL	573
KPGHKARPHIIRYKII	574
KIVIRIRIRLIRIRIRKIV	575
RIIVKIRLRIIKKIRLIKK	576
RIYVSKISIYIKKIRL	577
KIVIFTRIRLTSIRIRSIV	578
KPIHKARPTIIRYKMI	579
cyclicCKGFFALIPKIISSPLFKTLLSAVC	580
CKKGFFALIPKIISSPLFKTLLSAVC	581
CKKKGFFALIPKIISSPLFKTLLSAVC	582
CyclicCRIVIRIRIRLIRIRC	583
CyclicCKPGHKARPHIIRYKIIC	584
CyclicCRFAVKIRLRIIKKIRLIKKIRKRVIKC	585
KLLLKLLL KLLKC	586
KLLLKLLLKLLK	587
KLLLKLKLKLLKC	588
KLLLKLLLKLLK	589

Table 12—VIP-mimetic peptide sequences

Sequence/structure	SEQ ID NO:
HSDAVFYDNYTR LRKQMAVKKYLN SILN	590
Nle HSDAVFYDNYTR LRKQMAVKKYLN SILN	591
$X_1 X_1' X_1'' X_2$	592
X_3 S X_4 LN	593
NH CH CO KKYX5 NH CH CO X6 | | (CH2)m Z (CH2)n	594
KKYL	595
NSILN	596
KKYL	597
KKYA	598
AVKKYL	599
NSILN	600
KKYV	601
SILauN	602
KKYLNle	603
NSYLN	604
NSIYN	605
KKYLPPNSILN	606
LauKKYL	607
CapKKYL	608
KYL	NR
KKYNle	609
VKKYL	610
LNSILN	611
YLNSILN	612
KKYLN	613
KKYLNS	614
KKYLNSI	615
KKYLNSIL	616
KKYL	617
KKYDA	618
AVKKYL	619
NSILN	620
KKYV	621
SILauN	622
NSYLN	623
NSIYN	624
KKYLNle	625
KKYLPPNSILN	626
KKYL	627
KKYDA	628
AVKKYL	629
NSILN	630
KKYV	631
SILauN	632

LauKKYL	633
CapKKYL	634
KYL	NR
KYL	NR
KKYNle	635
VKKYL	636
LNSILN	637
YLNSILN	638
KKYLNle	639
KKYLN	640
KKYLNS	641
KKYLNSI	642
KKYLNSIL	643
KKKYLD	644
cyclicCKKYLC	645
CKKYLK | | S-CH_2-CO	646
KKYA	647
WWTDTGLW	648
WWTDDGLW	649
WWDTRGLWVWTI	650
FWGNDGIWLESG	651
DWDQFGLWRGAA	652
RWDDNGLWVVVL	653
SGMWSHYGIWMG	654
GGRWDQAGLWVA	655
KLWSEQGIWMGE	656
CWSMHGLWLC	657
GCWDNTGIWVPC	658
DWDTRGLWVY	659
SLWDENGAWI	660
KWDDRGLWMH	661
QAWNERGLWT	662
QWDTRGLWVA	663
WNVHGIWQE	664
SWDTRGLWVE	665
DWDTRGLWVA	666
SWGRDGLWIE	667
EWTDNGLWAL	668
SWDEKGLWSA	669
SWDSSGLWMD	670

Table 13—Mdm/hdm antagonist peptide sequences

Sequence/structure	SEQ ID NO:
TFSDLW	130
QETFSDLWKLLP	131
QPTFSDLWKLLP	132
QETFSDYWKLLP	133
QPTFSDYWKLLP	134
MPRFMDYWEGLN	135
VQNFIDYWTQQF	136
TGPAFTHYWATF	137
IDRAPTFRDHWFALV	138
PRPALVFADYWETLY	139
PAFSRFWSDLSAGAH	140
PAFSRFWSKLSAGAH	141
PXFXDYWXXL	142
QETFSDLWKLLP	143
QPTFSDLWKLLP	144
QETFSDYWKLLP	145
QPTFSDYWKLLP	146

Table 14—Calmodulin antagonist peptide sequences

Sequence/structure	SEQ ID NO:
SCVKWGKKEFCGS	164
SCWKYWGKECGS	165
SCYEWGKLRWCGS	166
SCLRWGKWSNCGS	167
SCWRWGKYQICGS	168
SCVSWGALKLCGS	169
SCIRWGQNTFCGS	170
SCWQWGNLKICGS	171
SCVRWGQLSICGS	172
LKKFNARRKLKGAILTTMLAK	173
RRWKKNFIAVSAANRFKK	174
RKWQKTGHAVRAIGRLSS	175
INLKALAALAKKIL	176
KIWSILAPLGTTLVKLVA	177
LKKLLKLLKKLLKL	178
LKWKKLLKLLKKLLKKLL	179
AEWPSLTEIKTLSHFSV	180
AEWPSPTRVISTTYFGS	181
AELAHWPPVKTVLRSFT	182
AEGSWLQLLNLMKQMNN	183
AEWPSLTEIK	184

Table 15—Mast cell antagonists/Mast cell protease inhibitor peptide sequences

Sequence/structure	SEQ ID NO:
SGSGVLKRPLPILPVTR	272
RWLSSRPLPPLPLPPRT	273
GSGSYDTLALPSLPLHPMSS	274
GSGSYDTRALPSLPLHPMSS	275
GSGSSGVTMYPKLPPHWSMA	276
GSGSSGVRMYPKLPPHWSMA	277
GSGSSSMRMVPTIPGSAKHG	278
RNR	NR
QT	NR
RQK	NR
NRQ	NR
RQK	NR
RNRQKT	436
RNRQ	437
RNRQK	438
NRQKT	439
RQKT	440

Table 16—SH3 antagonist peptide sequences

Sequence/structure	SEQ ID NO:
RPLPPLP	282
RELPPLP	283
SPLPPLP	284
GPLPPLP	285
RPLPIPP	286
RPLPIPP	287
RRLPPTP	288
RQLPPTP	289
RPLPSRP	290
RPLPTRP	291
SRLPPLP	292
RALPSPP	293
RRLPRTP	294
RPVPPIT	295
ILAPPVP	296
RPLPMLP	297
RPLPILP	298
RPLPSLP	299
RPLPSLP	300
RPLPMIP	301
RPLPLIP	302
RPLPPTP	303
RSLPPLP	304
RPQPPPP	305
RQLPIPP	306
XXXRPLPPLPXP	307
XXXRPLPPIPXX	308
XXXRPLPPLPXX	309
RXXRPLPPLPXP	310
RXXRPLPPLPPP	311
PPPYPPPPIPXX	312
PPPYPPPPVPXX	313
LXXRPLPXΨP	314
ΨXXRPLPXLP	315
PPXΘXPPPΨP	316
+PPΨPXKPXWL	317
RPXΨPΨR+SXP	318
PPVPPRPXXTL	319
ΨPΨLPΨK	320
+ΘDXPLPXLP	321

Table 17—Somatostatin or cortistatin mimetic peptide sequences

Sequence/structure	SEQ ID NO:
X^1-X^2-Asn-Phe-Phe-Trp-Lys-Thr-Phe-X^3-Ser-X^4	473
Asp Arg Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys	474
Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys	475
Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys	476
Asp Arg Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	477
Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	478
Cys Arg Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	479
Asp Arg Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	480
Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys	481
Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys	482
Asp Arg Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	483
Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	484
Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys	485
Asp Arg Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	486
Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	487
Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	488
Asp Arg Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	489
Met Pro Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	490
Cys Arg Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	491
Asp Arg Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	492
Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	493
Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys Lys	494
Asp Arg Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	495
Met Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	496
Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys	497

Table 18—UKR antagonist peptide sequences

Sequence/structure	SEQ ID NO:
AEPMPHSLNFSQYLWYT	196
AEHTYSSLWDTYSPLAF	197
AELDLWMRHYPLSFSNR	198
AESSLWTRYAWPSMPSY	199
AEWHPGLSFGSYLWSKT	200
AEPALLNWSFFFNPGLH	201
AEWSFYNLHLPEPQTIF	202
AEPLDLWSLYSLPPLAM	203
AEPTLWQLYQFPLRLSG	204
AEISFSELMWLRSTPAF	205
AELSEADLWTTWFGMGS	206
AESSLWRIFSPSALMMS	207
AESLPTLTSILWGKESV	208
AETLFMDLWHDKHILLT	209
AEILNFPLWHEPLWSTE	210
AESQTGTLNTLFWNTLR	211
AEPVYQYELDSYLRSYY	430
AELDLSTFYDIQYLLRT	431
AEFFKLGPNGYVYLHSA	432
FKLXXXGYVYL	433
AESTYHHLSLGYMYTLN	434
YHXLXXGYMYT	435

Table 19—Macrophage and/or T-cell inhibiting peptide sequences

Sequence/structure	SEQ ID NO:
Xaa-Yaa-Arg	NR
Arg-Yaa-Xaa	NR
Xaa-Arg-Yaa	NR
Yaa-Arg-Xaa	NR
Ala-Arg	NR
Arg-Arg	NR
Asn-Arg	NR
Asp-Arg	NR
Cys-Arg	NR
Gln-Arg	NR
Glu-Arg	NR
Gly-Arg	NR
His-arg	NR
Ile-Arg	NR
Leu-Arg	NR
Lys-Arg	NR
Met-Arg	NR
Phe-Arg	NR
Ser-Arg	NR
Thr-Arg	NR
Trp-Arg	NR
Tyr-Arg	NR
Val-Arg	NR
Ala-Glu-Arg	NR
Arg-Glu-Arg	NR
Asn-Glu-Arg	NR
Asp-Glu-Arg	NR
Cys-Glu-Arg	NR
Gln-Glu-Arg	NR
Glu-Glu-Arg	NR
Gly-Glu-Arg	NR
His-Glu-Arg	NR
Ile-Glu-Arg	NR
Leu-Glu-Arg	NR
Lys-Glu-Arg	NR
Met-Glu-Arg	NR
Phe-Glu-Arg	NR
Pro-Glu-Arg	NR
Ser-Glu-Arg	NR
Thr-Glu-Arg	NR
Trp-Glu-Arg	NR
Tyr-Glu-Arg	NR
Val-Glu-Arg	NR

Arg-Ala	NR
Arg-Asp	NR
Arg-Cys	NR
Arg-Gln	NR
Arg-Glu	NR
Arg-Gly	NR
Arg-His	NR
Arg-Ile	NR
Arg-Leu	NR
Arg-Lys	NR
Arg-Met	NR
Arg-Phe	NR
Arg-Pro	NR
Arg-Ser	NR
Arg-Thr	NR
Arg-Trp	NR
Arg-Tyr	NR
Arg-Val	NR
Arg-Glu-Ala	NR
Arg-Glu-Asn	NR
Arg-Glu-Asp	NR
Arg-Glu-Cys	NR
Arg-Glu-Gln	NR
Arg-Glu-Glu	NR
Arg-Glu-Gly	NR
Arg-Glu-His	NR
Arg-Glu-Ile	NR
Arg-Glu-Leu	NR
Arg-Glu-Lys	NR
Arg-Glu-Met	NR
Arg-Glu-Phe	NR
Arg-Glu-Pro	NR
Arg-Glu-Ser	NR
Arg-Glu-Thr	NR
Arg-Glu-Trp	NR
Arg-Glu-Tyr	NR
Arg-Glu-Val	NR
Ala-Arg-Glu	NR
Arg-Arg-Glu	NR
Asn-Arg-Glu	NR
Asp-Arg-Glu	NR
Cys-Arg-Glu	NR
Gln-Arg-Glu	NR
Glu-Arg-Glu	NR
Gly-Arg-Glu	NR
His-Arg-Glu	NR
Ile-Arg-Glu	NR
Leu-Arg-Glu	NR
Lys-Arg-Glu	NR
Met-Arg-Glu	NR

Phe-Arg-Glu	NR
Pro-Arg-Glu	NR
Ser-Arg-Glu	NR
Thr-Arg-Glu	NR
Trp-Arg-Glu	NR
Tyr-Arg-Glu	NR
Val-Arg-Glu	NR
Glu-Arg-Ala,	NR
Glu-Arg-Arg	NR
Glu-Arg-Asn	NR
Glu-Arg-Asp	NR
Glu-Arg-Cys	NR
Glu-Arg-Gln	NR
Glu-Arg-Gly	NR
Glu-Arg-His	NR
Glu-Arg-Ile	NR
Glu-Arg-Leu	NR
Glu-Arg-Lys	NR
Glu-Arg-Met	NR
Glu-Arg-Phe	NR
Glu-Arg-Pro	NR
Glu-Arg-Ser	NR
Glu-Arg-Thr	NR
Glu-Arg-Trp	NR
Glu-Arg-Tyr	NR
Glu-Arg-Val	NR

Table 20—Additional Exemplary Pharmacologically Active Peptides

Sequence/structure	SEQ ID NO:	Activity
VEPNCDIHVMWEWECFERL	1027	VEGF-antagonist
GERWCFDGPLTWVCGEES	1084	VEGF-antagonist
RGWVEICVADDNGMCVTEAQ	1085	VEGF-antagonist
GWDECDVARMWEWECFAGV	1086	VEGF- antagonist
GERWCFDGPRAWVCGWEI	501	VEGF- antagonist
EELWCFDGPRAWVCGYVK	502	VEGF- antagonist
RGWVEICAADDYGRCLTEAQ	1031	VEGF- antagonist
RGWVEICESDVWGRCL	1087	VEGF- antagonist
RGWVEICESDVWGRCL	1088	VEGF- antagonist
GGNECDIARMWEWECFERL	1089	VEGF- antagonist
RGWVEICAADDYGRCL	1090	VEGF-antagonist
CTTHWGFTLC	1028	MMP inhibitor
CLRSGXGC	1091	MMP inhibitor
CXXHWGFXXC	1092	MMP inhibitor
CXPXC	1093	MMP inhibitor
CRRHWGFEFC	1094	MMP inhibitor
STTHWGFTLS	1095	MMP inhibitor
CSLHWGFWWC	1096	CTLA4-mimetic
GFVCSGIFAVGVGRC	125	CTLA4-mimetic
APGVRLGCAVLGRYC	126	CTLA4-mimetic
LLGRMK	105	Antiviral (HBV)
ICVVQDWGHHRCTAGHMANLTSHASAI	127	C3b antagonist
ICVVQDWGHHRCT	128	C3b antagonist
CVVQDWGHHAC	129	C3b antagonist
STGGFDDVYDWARGVSSALTTTLVATR	185	Vinculin-binding
STGGFDDVYDWARRVSSALTTTLVATR	186	Vinculin-binding
SRGVNFSEWLYDMSAAMKEASNVFPSRRSR	187	Vinculin-binding
SSQNWDMEAGVEDLTAAMLGLLSTIHSSSR	188	Vinculin-binding
SSPSLYTQFLVNYESAATRIQDLLIASRPSR	189	Vinculin-binding
SSTGWVDLLGALQRAADATRTSIPPSLQNSR	190	Vinculin-binding
DVYTKKELIECARRVSEK	191	Vinculin-binding
EKGSYYPGSGIAQFHIDYNNVS	192	C4BP-binding
SGIAQFHIDYNNVSSAEGWHVN	193	C4BP-binding
LVTVEKGSYYPGSGIAQFHIDYNNVSSAEGWHVN	194	C4BP-binding
SGIAQFHIDYNNVS	195	C4BP-binding
LLGRMK	279	anti-HBV
ALLGRMKG	280	anti-HBV
LDPAFR	281	anti-HBV
CXXRGDC	322	Inhibition of platelet aggregation
RPLPPLP	323	Src antagonist
PPVPPR	324	Src antagonist
XFXDXWXXLXX	325	Anti-cancer (particularly for

		sarcomas)
KACRRLFGPVDSEQLSRDCD	326	p16-mimetic
RERWNFDFVTETPLEGDFAW	327	p16-mimetic
KRRQTSMTDFYHSKRRLIFS	328	p16-mimetic
TSMTDFYHSKRRLIFSKRKP	329	p16-mimetic
RRLIF	330	p16-mimetic
KRRQTSATDFYHSKRRLIFSRQIKIWFQNRRMKWKK	331	p16-mimetic
KRRLIFSKRQIKIWFQNRRMKWKK	332	p16-mimetic
Asn Gln Gly Arg His Phe Cys Gly Gly Ala Leu Ile His Ala Arg Phe Val Met Thr Ala Ala Ser Cys Phe Gln	498	CAP37 mimetic/LPS binding
Arg His Phe Cys Gly Gly Ala Leu Ile His Ala Arg Phe Val Met Thr Ala Ala Ser Cys	499	CAP37 mimetic/LPS binding
Gly Thr Arg Cys Gln Val Ala Gly Trp Gly Ser Gln Arg Ser Gly Gly Arg Leu Ser Arg Phe Pro Arg Phe Val Asn Val	500	CAP37 mimetic/LPS binding
WHWRHRIPLQLAAGR	1097	carbohydrate (GD1 alpha) mimetic
LKTPRV	1098	β2GPI Ab binding
NTLKTPRV	1099	β2GPI Ab binding
NTLKTPRVGGC	1100	β2GPI Ab binding
KDKATF	1101	β2GPI Ab binding
KDKATFGCHD	1102	β2GPI Ab binding
KDKATFGCHDGC	1103	β2GPI Ab binding
TLRVYK	1104	β2GPI Ab binding
ATLRVYKGG	1105	β2GPI Ab binding
CATLRVYKGG	1106	β2GPI Ab binding
INLKALAALAKKIL	1107	Membrane-transporting
GWT	NR	Membrane-transporting
GWTLNSAGYLLG	1108	Membrane-transporting
GWTLNSAGYLLGKINLKALAALAKKIL	1109	Membrane-transporting

The present invention is also particularly useful with peptides having activity in treatment of:

- cancer, wherein the peptide is a VEGF-mimetic or a VEGF receptor antagonist, a HER2 agonist or antagonist, a CD20 antagonist and the like;
- asthma, wherein the protein of interest is a CKR3 antagonist, an IL-5 receptor antagonist, and the like;
- thrombosis, wherein the protein of interest is a GPIIb antagonist, a GPIIIa antagonist, and the like;

- autoimmune diseases and other conditions involving immune modulation, wherein the protein of interest is an IL-2 receptor antagonist, a CD40 agonist or antagonist, a CD40L agonist or antagonist, a thymopoietin mimetic and the like.

Vehicles. This invention requires the presence of at least one vehicle (F^1, F^2) attached to a peptide through the N-terminus, C-terminus or a sidechain of one of the amino acid residues. Multiple vehicles may also be used; e.g., Fc's at each terminus or an Fc at a terminus and a PEG group at the other terminus or a sidechain.

An Fc domain is the preferred vehicle. The Fc domain may be fused to the N or C termini of the peptides or at both the N and C termini. For the TPO-mimetic peptides, molecules having the Fc domain fused to the N terminus of the peptide portion of the molecule are more bioactive than other such fusions, so fusion to the N terminus is preferred.

As noted above, Fc variants are suitable vehicles within the scope of this invention. A native Fc may be extensively modified to form an Fc variant in accordance with this invention, provided binding to the salvage receptor is maintained; see, for example WO 97/34631 and WO 96/32478. In such Fc variants, one may remove one or more sites of a native Fc that provide structural features or functional activity not required by the fusion molecules of this invention. One may remove these sites by, for example, substituting or deleting residues, inserting residues into the site, or truncating portions containing the site. The inserted or substituted residues may also be altered amino acids, such as peptidomimetics or D-amino acids. Fc variants may be desirable for a number of reasons, several of which are described below. Exemplary Fc variants include molecules and sequences in which:

1. Sites involved in disulfide bond formation are removed. Such removal may avoid reaction with other cysteine-containing proteins present in

the host cell used to produce the molecules of the invention. For this purpose, the cysteine-containing segment at the N-terminus may be truncated or cysteine residues may be deleted or substituted with other amino acids (e.g., alanyl, seryl). In particular, one may truncate the N-terminal 20-amino acid segment of SEQ ID NO: 2 or delete or substitute the cysteine residues at positions 7 and 10 of SEQ ID NO: 2. Even when cysteine residues are removed, the single chain Fc domains can still form a dimeric Fc domain that is held together non-covalently.

2. A native Fc is modified to make it more compatible with a selected host cell. For example, one may remove the PA sequence near the N-terminus of a typical native Fc, which may be recognized by a digestive enzyme in E. coli such as proline iminopeptidase. One may also add an N-terminal methionine residue, especially when the molecule is expressed recombinantly in a bacterial cell such as E. coli. The Fc domain of SEQ ID NO: 2 (Figure 4) is one such Fc variant.
3. A portion of the N-terminus of a native Fc is removed to prevent N-terminal heterogeneity when expressed in a selected host cell. For this purpose, one may delete any of the first 20 amino acid residues at the N-terminus, particularly those at positions 1, 2, 3, 4 and 5.
4. One or more glycosylation sites are removed. Residues that are typically glycosylated (e.g., asparagine) may confer cytolytic response. Such residues may be deleted or substituted with unglycosylated residues (e.g., alanine).
5. Sites involved in interaction with complement, such as the C1q binding site, are removed. For example, one may delete or substitute the EKK sequence of human IgG1. Complement recruitment may not be advantageous for the molecules of this invention and so may be avoided with such an Fc variant.

6. Sites are removed that affect binding to Fc receptors other than a salvage receptor. A native Fc may have sites for interaction with certain white blood cells that are not required for the fusion molecules of the present invention and so may be removed.
7. The ADCC site is removed. ADCC sites are known in the art; see, for example, Molec. Immunol. 29 (5): 633-9 (1992) with regard to ADCC sites in IgG1. These sites, as well, are not required for the fusion molecules of the present invention and so may be removed.
8. When the native Fc is derived from a non-human antibody, the native Fc may be humanized. Typically, to humanize a native Fc, one will substitute selected residues in the non-human native Fc with residues that are normally found in human native Fc. Techniques for antibody humanization are well known in the art.

Preferred Fc variants include the following. In SEQ ID NO: 2 (Figure 4) the leucine at position 15 may be substituted with glutamate; the glutamate at position 99, with alanine; and the lysines at positions 101 and 103, with alanines. In addition, one or more tyrosine residues can be replaced by phenyalanine residues.

An alternative vehicle would be a protein, polypeptide, peptide, antibody, antibody fragment, , or small molecule (e.g., a peptidomimetic compound) capable of binding to a salvage receptor. For example, one could use as a vehicle a polypeptide as described in U.S. Pat. No. 5,739,277, issued April 14, 1998 to Presta et al. Peptides could also be selected by phage display for binding to the FcRn salvage receptor. Such salvage receptor-binding compounds are also included within the meaning of "vehicle" and are within the scope of this invention. Such vehicles should be selected for increased half-life (e.g., by avoiding sequences recognized by proteases) and decreased immunogenicity (e.g., by favoring non-immunogenic sequences, as discovered in antibody humanization).

As noted above, polymer vehicles may also be used for F^1 and F^2. Various means for attaching chemical moieties useful as vehicles are currently available, see, e.g., Patent Cooperation Treaty ("PCT") International Publication No. WO 96/11953, entitled "N-Terminally Chemically Modified Protein Compositions and Methods," herein incorporated by reference in its entirety. This PCT publication discloses, among other things, the selective attachment of water soluble polymers to the N-terminus of proteins.

A preferred polymer vehicle is polyethylene glycol (PEG). The PEG group may be of any convenient molecular weight and may be linear or branched. The average molecular weight of the PEG will preferably range from about 2 kiloDalton ("kD") to about 100 kDa, more preferably from about 5 kDa to about 50 kDa, most preferably from about 5 kDa to about 10 kDa. The PEG groups will generally be attached to the compounds of the invention via acylation or reductive alkylation through a reactive group on the PEG moiety (e.g., an aldehyde, amino, thiol, or ester group) to a reactive group on the inventive compound (e.g., an aldehyde, amino, or ester group).

A useful strategy for the PEGylation of synthetic peptides consists of combining, through forming a conjugate linkage in solution, a peptide and a PEG moiety, each bearing a special functionality that is mutually reactive toward the other. The peptides can be easily prepared with conventional solid phase synthesis (see, for example, Figures 5 and 6 and the accompanying text herein). The peptides are "preactivated" with an appropriate functional group at a specific site. The precursors are purified and fully characterized prior to reacting with the PEG moiety. Ligation of the peptide with PEG usually takes place in aqueous phase and can be easily monitored by reverse phase analytical HPLC. The PEGylated peptides can be easily purified by preparative HPLC and characterized by

analytical HPLC, amino acid analysis and laser desorption mass spectrometry.

Polysaccharide polymers are another type of water soluble polymer which may be used for protein modification. Dextrans are polysaccharide polymers comprised of individual subunits of glucose predominantly linked by α1-6 linkages. The dextran itself is available in many molecular weight ranges, and is readily available in molecular weights from about 1 kD to about 70 kD. Dextran is a suitable water soluble polymer for use in the present invention as a vehicle by itself or in combination with another vehicle (e.g., Fc). See, for example, WO 96/11953 and WO 96/05309. The use of dextran conjugated to therapeutic or diagnostic immunoglobulins has been reported; see, for example, European Patent Publication No. 0 315 456, which is hereby incorporated by reference. Dextran of about 1 kD to about 20 kD is preferred when dextran is used as a vehicle in accordance with the present invention.

Linkers. Any "linker" group is optional. When present, its chemical structure is not critical, since it serves primarily as a spacer. The linker is preferably made up of amino acids linked together by peptide bonds. Thus, in preferred embodiments, the linker is made up of from 1 to 20 amino acids linked by peptide bonds, wherein the amino acids are selected from the 20 naturally occurring amino acids. Some of these amino acids may be glycosylated, as is well understood by those in the art. In a more preferred embodiment, the 1 to 20 amino acids are selected from glycine, alanine, proline, asparagine, glutamine, and lysine. Even more preferably, a linker is made up of a majority of amino acids that are sterically unhindered, such as glycine and alanine. Thus, preferred linkers are polyglycines (particularly $(Gly)_4$, $(Gly)_5$), poly(Gly-Ala), and polyalanines. Other specific examples of linkers are:

$(Gly)_3Lys(Gly)_4$ (SEQ ID NO: 333);

$(Gly)_3AsnGlySer(Gly)_2$ (SEQ ID NO: 334);

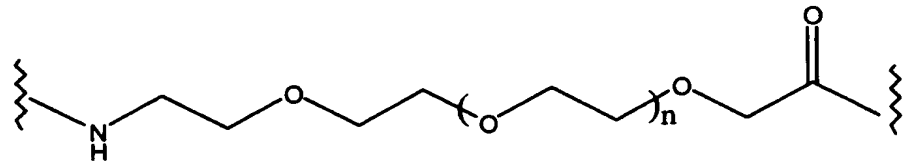
$(Gly)_3Cys(Gly)_4$ (SEQ ID NO: 335); and

GlyProAsnGlyGly (SEQ ID NO: 336).

To explain the above nomenclature, for example, $(Gly)_3Lys(Gly)_4$ means Gly-Gly-Gly-Lys-Gly-Gly-Gly-Gly. Combinations of Gly and Ala are also preferred. The linkers shown here are exemplary; linkers within the scope of this invention may be much longer and may include other residues.

Non-peptide linkers are also possible. For example, alkyl linkers such as $-NH-(CH_2)_s-C(O)-$, wherein s = 2-20 could be used. These alkyl linkers may further be substituted by any non-sterically hindering group such as lower alkyl (e.g., C_1-C_6) lower acyl, halogen (e.g., Cl, Br), CN, NH_2, phenyl, etc. An exemplary non-peptide linker is a PEG linker,

VI



wherein n is such that the linker has a molecular weight of 100 to 5000 kD, preferably 100 to 500 kD. The peptide linkers may be altered to form derivatives in the same manner as described above.

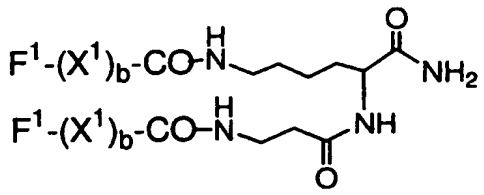
Derivatives. The inventors also contemplate derivatizing the peptide and/or vehicle portion of the compounds. Such derivatives may improve the solubility, absorption, biological half life, and the like of the compounds. The moieties may alternatively eliminate or attenuate any undesirable side-effect of the compounds and the like. Exemplary derivatives include compounds in which:

1. The compound or some portion thereof is cyclic. For example, the peptide portion may be modified to contain two or more Cys residues (e.g., in the linker), which could cyclize by disulfide bond formation.

For citations to references on preparation of cyclized derivatives, see Table 2.

2. The compound is cross-linked or is rendered capable of cross-linking between molecules. For example, the peptide portion may be modified to contain one Cys residue and thereby be able to form an intermolecular disulfide bond with a like molecule. The compound may also be cross-linked through its C-terminus, as in the molecule shown below.

VII



3.

4. One or more peptidyl [-C(O)NR-] linkages (bonds) is replaced by a non-peptidyl linkage. Exemplary non-peptidyl linkages are -CH_2-carbamate [-CH_2-OC(O)NR-], phosphonate , -CH_2-sulfonamide [-CH_2-$S(O)_2$NR-], urea [-NHC(O)NH-], -CH_2-secondary amine, and alkylated peptide [-C(O)NR^6- wherein R^6 is lower alkyl].

5. The N-terminus is derivatized. Typically, the N-terminus may be acylated or modified to a substituted amine. Exemplary N-terminal derivative groups include -NRR^1 (other than -NH_2), -$NRC(O)R^1$, -$NRC(O)OR^1$, -$NRS(O)_2R^1$, -$NHC(O)NHR^1$, succinimide, or benzyloxycarbonyl-NH- (CBZ-NH-), wherein R and R^1 are each independently hydrogen or lower alkyl and wherein the phenyl ring may be substituted with 1 to 3 substituents selected from the group consisting of C_1-C_4 alkyl, C_1-C_4 alkoxy, chloro, and bromo.

6. The free C-terminus is derivatized. Typically, the C-terminus is esterified or amidated. For example, one may use methods described in the art to add $(NH\text{-}CH_2\text{-}CH_2\text{-}NH_2)_2$ to compounds of this invention

having any of SEQ ID NOS: 504 to 508 at the C-terminus. Likewise, one may use methods described in the art to add -NH_2 to compounds of this invention having any of SEQ ID NOS: 924 to 955, 963 to 972, 1005 to 1013, or 1018 to 1023 at the C-terminus. Exemplary C-terminal derivative groups include, for example, -$C(O)R^2$ wherein R^2 is lower alkoxy or -NR^3R^4 wherein R^3 and R^4 are independently hydrogen or C_1-C_8 alkyl (preferably C_1-C_4 alkyl).

7. A disulfide bond is replaced with another, preferably more stable, cross-linking moiety (e.g., an alkylene). See, e.g., Bhatnagar et al. (1996), J. Med. Chem. 39: 3814-9; Alberts et al. (1993) Thirteenth Am. Pep. Symp., 357-9.
8. One or more individual amino acid residues is modified. Various derivatizing agents are known to react specifically with selected sidechains or terminal residues, as described in detail below.

Lysinyl residues and amino terminal residues may be reacted with succinic or other carboxylic acid anhydrides, which reverse the charge of the lysinyl residues. Other suitable reagents for derivatizing alpha-amino-containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; chloroborohydride; trinitrobenzenesulfonic acid; O-methylisourea; 2,4 pentanedione; and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues may be modified by reaction with any one or combination of several conventional reagents, including phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and ninhydrin. Derivatization of arginyl residues requires that the reaction be performed in alkaline conditions because of the high pKa of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine epsilon-amino group.

Specific modification of tyrosyl residues has been studied extensively, with particular interest in introducing spectral labels into tyrosyl residues by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizole and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl sidechain groups (aspartyl or glutamyl) may be selectively modified by reaction with carbodiimides (R'-N=C=N-R') such as 1-cyclohexyl-3-(2-morpholinyl-(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residues may be converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues may be deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Cysteinyl residues can be replaced by amino acid residues or other moieties either to eliminate disulfide bonding or, conversely, to stabilize cross-linking. See, e.g., Bhatnagar et al. (1996), J. Med. Chem. 39: 3814-9.

Derivatization with bifunctional agents is useful for cross-linking the peptides or their functional derivatives to a water-insoluble support matrix or to other macromolecular vehicles. Commonly used cross-linking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate yield photoactivatable intermediates that are capable of forming crosslinks in the presence of light. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates

and the reactive substrates described in U.S. Pat. Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440 are employed for protein immobilization.

Carbohydrate (oligosaccharide) groups may conveniently be attached to sites that are known to be glycosylation sites in proteins. Generally, O-linked oligosaccharides are attached to serine (Ser) or threonine (Thr) residues while N-linked oligosaccharides are attached to asparagine (Asn) residues when they are part of the sequence Asn-X-Ser/Thr, where X can be any amino acid except proline. X is preferably one of the 19 naturally occurring amino acids other than proline. The structures of N-linked and O-linked oligosaccharides and the sugar residues found in each type are different. One type of sugar that is commonly found on both is N-acetylneuraminic acid (referred to as sialic acid). Sialic acid is usually the terminal residue of both N-linked and O-linked oligosaccharides and, by virtue of its negative charge, may confer acidic properties to the glycosylated compound. Such site(s) may be incorporated in the linker of the compounds of this invention and are preferably glycosylated by a cell during recombinant production of the polypeptide compounds (e.g., in mammalian cells such as CHO, BHK, COS). However, such sites may further be glycosylated by synthetic or semi-synthetic procedures known in the art.

Other possible modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, oxidation of the sulfur atom in Cys, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains. Creighton, Proteins: Structure and Molecule Properties (W. H. Freeman & Co., San Francisco), pp. 79-86 (1983).

Compounds of the present invention may be changed at the DNA level, as well. The DNA sequence of any portion of the compound may be

changed to codons more compatible with the chosen host cell. For E. coli, which is the preferred host cell, optimized codons are known in the art. Codons may be substituted to eliminate restriction sites or to include silent restriction sites, which may aid in processing of the DNA in the selected host cell. The vehicle, linker and peptide DNA sequences may be modified to include any of the foregoing sequence changes.

Methods of Making

The compounds of this invention largely may be made in transformed host cells using recombinant DNA techniques. To do so, a recombinant DNA molecule coding for the peptide is prepared. Methods of preparing such DNA molecules are well known in the art. For instance, sequences coding for the peptides could be excised from DNA using suitable restriction enzymes. Alternatively, the DNA molecule could be synthesized using chemical synthesis techniques, such as the phosphoramidate method. Also, a combination of these techniques could be used.

The invention also includes a vector capable of expressing the peptides in an appropriate host. The vector comprises the DNA molecule that codes for the peptides operatively linked to appropriate expression control sequences. Methods of effecting this operative linking, either before or after the DNA molecule is inserted into the vector, are well known. Expression control sequences include promoters, activators, enhancers, operators, ribosomal binding sites, start signals, stop signals, cap signals, polyadenylation signals, and other signals involved with the control of transcription or translation.

The resulting vector having the DNA molecule thereon is used to transform an appropriate host. This transformation may be performed using methods well known in the art.

Any of a large number of available and well-known host cells may be used in the practice of this invention. The selection of a particular host is dependent upon a number of factors recognized by the art. These include, for example, compatibility with the chosen expression vector, toxicity of the peptides encoded by the DNA molecule, rate of transformation, ease of recovery of the peptides, expression characteristics, bio-safety and costs. A balance of these factors must be struck with the understanding that not all hosts may be equally effective for the expression of a particular DNA sequence. Within these general guidelines, useful microbial hosts include bacteria (such as E. coli sp.), yeast (such as Saccharomyces sp.) and other fungi, insects, plants, mammalian (including human) cells in culture, or other hosts known in the art.

Next, the transformed host is cultured and purified. Host cells may be cultured under conventional fermentation conditions so that the desired compounds are expressed. Such fermentation conditions are well known in the art. Finally, the peptides are purified from culture by methods well known in the art.

The compounds may also be made by synthetic methods. For example, solid phase synthesis techniques may be used. Suitable techniques are well known in the art, and include those described in Merrifield (1973), Chem. Polypeptides, pp. 335-61 (Katsoyannis and Panayotis eds.); Merrifield (1963), J. Am. Chem. Soc. 85: 2149; Davis et al. (1985), Biochem. Intl. 10: 394-414; Stewart and Young (1969), Solid Phase Peptide Synthesis; U.S. Pat. No. 3,941,763; Finn et al. (1976), The Proteins (3rd ed.) 2: 105-253; and Erickson et al. (1976), The Proteins (3rd ed.) 2: 257-527. Solid phase synthesis is the preferred technique of making individual peptides since it is the most cost-effective method of making small peptides.

Compounds that contain derivatized peptides or which contain non-peptide groups may be synthesized by well-known organic chemistry techniques.

Uses of the Compounds

In general. The compounds of this invention have pharmacologic activity resulting from their ability to bind to proteins of interest as agonists, mimetics or antagonists of the native ligands of such proteins of interest. The utility of specific compounds is shown in Table 2. The activity of these compounds can be measured by assays known in the art. For the TPO-mimetic and EPO-mimetic compounds, in vivo assays are further described in the Examples section herein.

In addition to therapeutic uses, the compounds of the present invention are useful in diagnosing diseases characterized by dysfunction of their associated protein of interest. In one embodiment, a method of detecting in a biological sample a protein of interest (e.g., a receptor) that is capable of being activated comprising the steps of: (a) contacting the sample with a compound of this invention; and (b) detecting activation of the protein of interest by the compound. The biological samples include tissue specimens, intact cells, or extracts thereof. The compounds of this invention may be used as part of a diagnostic kit to detect the presence of their associated proteins of interest in a biological sample. Such kits employ the compounds of the invention having an attached label to allow for detection. The compounds are useful for identifying normal or abnormal proteins of interest. For the EPO-mimetic compounds, for example, presence of abnormal protein of interest in a biological sample may be indicative of such disorders as Diamond Blackfan anemia, where it is believed that the EPO receptor is dysfunctional.

Therapeutic uses of EPO-mimetic compounds. The EPO-mimetic compounds of the invention are useful for treating disorders characterized by low red blood cell levels. Included in the invention are methods of modulating the endogenous activity of an EPO receptor in a mammal, preferably methods of increasing the activity of an EPO receptor. In

general, any condition treatable by erythropoietin, such as anemia, may also be treated by the EPO-mimetic compounds of the invention. These compounds are administered by an amount and route of delivery that is appropriate for the nature and severity of the condition being treated and may be ascertained by one skilled in the art. Preferably, administration is by injection, either subcutaneous, intramuscular, or intravenous.

Therapeutic uses of TPO-mimetic compounds. For the TPO-mimetic compounds, one can utilize such standard assays as those described in WO95/26746 entitled "Compositions and Methods for Stimulating Megakaryocyte Growth and Differentiation". In vivo assays also appear in the Examples hereinafter.

The conditions to be treated are generally those that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency (e.g., because of planned surgery or platelet donation). Such conditions will usually be the result of a deficiency (temporary or permanent) of active Mpl ligand in vivo. The generic term for platelet deficiency is thrombocytopenia, and hence the methods and compositions of the present invention are generally available for treating thrombocytopenia in patients in need thereof.

Thrombocytopenia (platelet deficiencies) may be present for various reasons, including chemotherapy and other therapy with a variety of drugs, radiation therapy, surgery, accidental blood loss, and other specific disease conditions. Exemplary specific disease conditions that involve thrombocytopenia and may be treated in accordance with this invention are: aplastic anemia, idiopathic thrombocytopenia, metastatic tumors which result in thrombocytopenia, systemic lupus erythematosus, splenomegaly, Fanconi's syndrome, vitamin B12 deficiency, folic acid deficiency, May-Hegglin anomaly, Wiskott-Aldrich syndrome, and paroxysmal nocturnal hemoglobinuria. Also, certain treatments for AIDS

result in thrombocytopenia (e.g., AZT). Certain wound healing disorders might also benefit from an increase in platelet numbers.

With regard to anticipated platelet deficiencies, e.g., due to future surgery, a compound of the present invention could be administered several days to several hours prior to the need for platelets. With regard to acute situations, e.g., accidental and massive blood loss, a compound of this invention could be administered along with blood or purified platelets.

The TPO-mimetic compounds of this invention may also be useful in stimulating certain cell types other than megakaryocytes if such cells are found to express Mpl receptor. Conditions associated with such cells that express the Mpl receptor, which are responsive to stimulation by the Mpl ligand, are also within the scope of this invention.

The TPO-mimetic compounds of this invention may be used in any situation in which production of platelets or platelet precursor cells is desired, or in which stimulation of the c-Mpl receptor is desired. Thus, for example, the compounds of this invention may be used to treat any condition in a mammal wherein there is a need of platelets, megakaryocytes, and the like. Such conditions are described in detail in the following exemplary sources: WO95/26746; WO95/21919; WO95/18858; WO95/21920 and are incorporated herein.

The TPO-mimetic compounds of this invention may also be useful in maintaining the viability or storage life of platelets and/or megakaryocytes and related cells. Accordingly, it could be useful to include an effective amount of one or more such compounds in a composition containing such cells.

The therapeutic methods, compositions and compounds of the present invention may also be employed, alone or in combination with other cytokines, soluble Mpl receptor, hematopoietic factors, interleukins, growth factors or antibodies in the treatment of disease states

characterized by other symptoms as well as platelet deficiencies. It is anticipated that the inventive compound will prove useful in treating some forms of thrombocytopenia in combination with general stimulators of hematopoiesis, such as IL-3 or GM-CSF. Other megakaryocytic stimulatory factors, i.e., meg-CSF, stem cell factor (SCF), leukemia inhibitory factor (LIF), oncostatin M (OSM), or other molecules with megakaryocyte stimulating activity may also be employed with Mpl ligand. Additional exemplary cytokines or hematopoietic factors for such co-administration include IL-1 alpha, IL-1 beta, IL-2, IL-3, IL-4, IL-5, IL-6, IL-11, colony stimulating factor-1 (CSF-1), SCF, GM-CSF, granulocyte colony stimulating factor (G-CSF), EPO, interferon-alpha (IFN-alpha), consensus interferon, IFN-beta, or IFN-gamma. It may further be useful to administer, either simultaneously or sequentially, an effective amount of a soluble mammalian Mpl receptor, which appears to have an effect of causing megakaryocytes to fragment into platelets once the megakaryocytes have reached mature form. Thus, administration of an inventive compound (to enhance the number of mature megakaryocytes) followed by administration of the soluble Mpl receptor (to inactivate the ligand and allow the mature megakaryocytes to produce platelets) is expected to be a particularly effective means of stimulating platelet production. The dosage recited above would be adjusted to compensate for such additional components in the therapeutic composition. Progress of the treated patient can be monitored by conventional methods.

In cases where the inventive compounds are added to compositions of platelets and/or megakaryocytes and related cells, the amount to be included will generally be ascertained experimentally by techniques and assays known in the art. An exemplary range of amounts is 0.1 μg—1 mg inventive compound per 10^6 cells.

Pharmaceutical Compositions

In General. The present invention also provides methods of using pharmaceutical compositions of the inventive compounds. Such pharmaceutical compositions may be for administration for injection, or for oral, pulmonary, nasal, transdermal or other forms of administration. In general, the invention encompasses pharmaceutical compositions comprising effective amounts of a compound of the invention together with pharmaceutically acceptable diluents, preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. Such compositions include diluents of various buffer content (*e.g.*, Tris-HCl, acetate, phosphate), pH and ionic strength; additives such as detergents and solubilizing agents (*e.g.*, Tween 80, Polysorbate 80), anti-oxidants (*e.g.*, ascorbic acid, sodium metabisulfite), preservatives (*e.g.*, Thimersol, benzyl alcohol) and bulking substances (*e.g.*, lactose, mannitol); incorporation of the material into particulate preparations of polymeric compounds such as polylactic acid, polyglycolic acid, etc. or into liposomes. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation. Such compositions may influence the physical state, stability, rate of in vivo release, and rate of in vivo clearance of the present proteins and derivatives. See, e.g., Remington's Pharmaceutical Sciences, 18th Ed. (1990, Mack Publishing Co., Easton, PA 18042) pages 1435-1712 which are herein incorporated by reference. The compositions may be prepared in liquid form, or may be in dried powder, such as lyophilized form. Implantable sustained release formulations are also contemplated, as are transdermal formulations.

Oral dosage forms. Contemplated for use herein are oral solid dosage forms, which are described generally in Chapter 89 of Remington's Pharmaceutical Sciences (1990), 18th Ed., Mack Publishing Co. Easton PA 18042, which is herein incorporated by reference. Solid dosage forms include tablets, capsules, pills, troches or lozenges, cachets or pellets. Also,

liposomal or proteinoid encapsulation may be used to formulate the present compositions (as, for example, proteinoid microspheres reported in U.S. Patent No. 4,925,673). Liposomal encapsulation may be used and the liposomes may be derivatized with various polymers (e.g., U.S. Patent No. 5,013,556). A description of possible solid dosage forms for the therapeutic is given in Chapter 10 of Marshall, K., Modern Pharmaceutics (1979), edited by G. S. Banker and C. T. Rhodes, herein incorporated by reference. In general, the formulation will include the inventive compound, and inert ingredients which allow for protection against the stomach environment, and release of the biologically active material in the intestine.

Also specifically contemplated are oral dosage forms of the above inventive compounds. If necessary, the compounds may be chemically modified so that oral delivery is efficacious. Generally, the chemical modification contemplated is the attachment of at least one moiety to the compound molecule itself, where said moiety permits (a) inhibition of proteolysis; and (b) uptake into the blood stream from the stomach or intestine. Also desired is the increase in overall stability of the compound and increase in circulation time in the body. Moieties useful as covalently attached vehicles in this invention may also be used for this purpose. Examples of such moieties include: PEG, copolymers of ethylene glycol and propylene glycol, carboxymethyl cellulose, dextran, polyvinyl alcohol, polyvinyl pyrrolidone and polyproline. See, for example, Abuchowski and Davis, Soluble Polymer-Enzyme Adducts, Enzymes as Drugs (1981), Hocenberg and Roberts, eds., Wiley-Interscience, New York, NY, , pp 367-83; Newmark, et al. (1982), J. Appl. Biochem. 4:185-9. Other polymers that could be used are poly-1,3-dioxolane and poly-1,3,6-tioxocane. Preferred for pharmaceutical usage, as indicated above, are PEG moieties.

For oral delivery dosage forms, it is also possible to use a salt of a modified aliphatic amino acid, such as sodium N-(8-[2-hydroxybenzoyl] amino) caprylate (SNAC), as a carrier to enhance absorption of the therapeutic compounds of this invention. The clinical efficacy of a heparin formulation using SNAC has been demonstrated in a Phase II trial conducted by Emisphere Technologies. See US Patent No. 5,792,451, "Oral drug delivery composition and methods".

The compounds of this invention can be included in the formulation as fine multiparticulates in the form of granules or pellets of particle size about 1 mm. The formulation of the material for capsule administration could also be as a powder, lightly compressed plugs or even as tablets. The therapeutic could be prepared by compression.

Colorants and flavoring agents may all be included. For example, the protein (or derivative) may be formulated (such as by liposome or microsphere encapsulation) and then further contained within an edible product, such as a refrigerated beverage containing colorants and flavoring agents.

One may dilute or increase the volume of the compound of the invention with an inert material. These diluents could include carbohydrates, especially mannitol, α-lactose, anhydrous lactose, cellulose, sucrose, modified dextrans and starch. Certain inorganic salts may also be used as fillers including calcium triphosphate, magnesium carbonate and sodium chloride. Some commercially available diluents are Fast-Flo, Emdex, STA-Rx 1500, Emcompress and Avicell.

Disintegrants may be included in the formulation of the therapeutic into a solid dosage form. Materials used as disintegrants include but are not limited to starch including the commercial disintegrant based on starch, Explotab. Sodium starch glycolate, Amberlite, sodium carboxymethylcellulose, ultramylopectin, sodium alginate, gelatin, orange

peel, acid carboxymethyl cellulose, natural sponge and bentonite may all be used. Another form of the disintegrants are the insoluble cationic exchange resins. Powdered gums may be used as disintegrants and as binders and these can include powdered gums such as agar, Karaya or tragacanth. Alginic acid and its sodium salt are also useful as disintegrants.

Binders may be used to hold the therapeutic agent together to form a hard tablet and include materials from natural products such as acacia, tragacanth, starch and gelatin. Others include methyl cellulose (MC), ethyl cellulose (EC) and carboxymethyl cellulose (CMC). Polyvinyl pyrrolidone (PVP) and hydroxypropylmethyl cellulose (HPMC) could both be used in alcoholic solutions to granulate the therapeutic.

An antifrictional agent may be included in the formulation of the therapeutic to prevent sticking during the formulation process. Lubricants may be used as a layer between the therapeutic and the die wall, and these can include but are not limited to; stearic acid including its magnesium and calcium salts, polytetrafluoroethylene (PTFE), liquid paraffin, vegetable oils and waxes. Soluble lubricants may also be used such as sodium lauryl sulfate, magnesium lauryl sulfate, polyethylene glycol of various molecular weights, Carbowax 4000 and 6000.

Glidants that might improve the flow properties of the drug during formulation and to aid rearrangement during compression might be added. The glidants may include starch, talc, pyrogenic silica and hydrated silicoaluminate.

To aid dissolution of the compound of this invention into the aqueous environment a surfactant might be added as a wetting agent. Surfactants may include anionic detergents such as sodium lauryl sulfate, dioctyl sodium sulfosuccinate and dioctyl sodium sulfonate. Cationic detergents might be used and could include benzalkonium chloride or

benzethonium chloride. The list of potential nonionic detergents that could be included in the formulation as surfactants are lauromacrogol 400, polyoxyl 40 stearate, polyoxyethylene hydrogenated castor oil 10, 50 and 60, glycerol monostearate, polysorbate 40, 60, 65 and 80, sucrose fatty acid ester, methyl cellulose and carboxymethyl cellulose. These surfactants could be present in the formulation of the protein or derivative either alone or as a mixture in different ratios.

Additives may also be included in the formulation to enhance uptake of the compound. Additives potentially having this property are for instance the fatty acids oleic acid, linoleic acid and linolenic acid.

Controlled release formulation may be desirable. The compound of this invention could be incorporated into an inert matrix which permits release by either diffusion or leaching mechanisms e.g., gums. Slowly degenerating matrices may also be incorporated into the formulation, e.g., alginates, polysaccharides. Another form of a controlled release of the compounds of this invention is by a method based on the Oros therapeutic system (Alza Corp.), i.e., the drug is enclosed in a semipermeable membrane which allows water to enter and push drug out through a single small opening due to osmotic effects. Some enteric coatings also have a delayed release effect.

Other coatings may be used for the formulation. These include a variety of sugars which could be applied in a coating pan. The therapeutic agent could also be given in a film coated tablet and the materials used in this instance are divided into 2 groups. The first are the nonenteric materials and include methyl cellulose, ethyl cellulose, hydroxyethyl cellulose, methylhydroxy-ethyl cellulose, hydroxypropyl cellulose, hydroxypropyl-methyl cellulose, sodium carboxy-methyl cellulose, providone and the polyethylene glycols. The second group consists of the enteric materials that are commonly esters of phthalic acid.

A mix of materials might be used to provide the optimum film coating. Film coating may be carried out in a pan coater or in a fluidized bed or by compression coating.

Pulmonary delivery forms. Also contemplated herein is pulmonary delivery of the present protein (or derivatives thereof). The protein (or derivative) is delivered to the lungs of a mammal while inhaling and traverses across the lung epithelial lining to the blood stream. (Other reports of this include Adjei et al., Pharma. Res. (1990) 7: 565-9; Adjei et al. (1990), Internatl. J. Pharmaceutics 63: 135-44 (leuprolide acetate); Braquet et al. (1989), J. Cardiovasc. Pharmacol. 13 (suppl.5): s.143-146 (endothelin-1); Hubbard et al. (1989), Annals Int. Med. 3: 206-12 (α1-antitrypsin); Smith et al. (1989), J. Clin. Invest. 84: 1145-6 (α1-proteinase); Oswein et al. (March 1990), "Aerosolization of Proteins", Proc. Symp. Resp. Drug Delivery II, Keystone, Colorado (recombinant human growth hormone); Debs et al. (1988), J. Immunol. 140: 3482-8 (interferon-γ and tumor necrosis factor α) and Platz et al., U.S. Patent No. 5,284,656 (granulocyte colony stimulating factor).

Contemplated for use in the practice of this invention are a wide range of mechanical devices designed for pulmonary delivery of therapeutic products, including but not limited to nebulizers, metered dose inhalers, and powder inhalers, all of which are familiar to those skilled in the art. Some specific examples of commercially available devices suitable for the practice of this invention are the Ultravent nebulizer, manufactured by Mallinckrodt, Inc., St. Louis, Missouri; the Acorn II nebulizer, manufactured by Marquest Medical Products, Englewood, Colorado; the Ventolin metered dose inhaler, manufactured by Glaxo Inc., Research Triangle Park, North Carolina; and the Spinhaler powder inhaler, manufactured by Fisons Corp., Bedford, Massachusetts.

All such devices require the use of formulations suitable for the dispensing of the inventive compound. Typically, each formulation is specific to the type of device employed and may involve the use of an appropriate propellant material, in addition to diluents, adjuvants and/or carriers useful in therapy.

The inventive compound should most advantageously be prepared in particulate form with an average particle size of less than 10 μm (or microns), most preferably 0.5 to 5 μm, for most effective delivery to the distal lung.

Pharmaceutically acceptable carriers include carbohydrates such as trehalose, mannitol, xylitol, sucrose, lactose, and sorbitol. Other ingredients for use in formulations may include DPPC, DOPE, DSPC and DOPC. Natural or synthetic surfactants may be used. PEG may be used (even apart from its use in derivatizing the protein or analog). Dextrans, such as cyclodextran, may be used. Bile salts and other related enhancers may be used. Cellulose and cellulose derivatives may be used. Amino acids may be used, such as use in a buffer formulation.

Also, the use of liposomes, microcapsules or microspheres, inclusion complexes, or other types of carriers is contemplated.

Formulations suitable for use with a nebulizer, either jet or ultrasonic, will typically comprise the inventive compound dissolved in water at a concentration of about 0.1 to 25 mg of biologically active protein per mL of solution. The formulation may also include a buffer and a simple sugar (e.g., for protein stabilization and regulation of osmotic pressure). The nebulizer formulation may also contain a surfactant, to reduce or prevent surface induced aggregation of the protein caused by atomization of the solution in forming the aerosol.

Formulations for use with a metered-dose inhaler device will generally comprise a finely divided powder containing the inventive

compound suspended in a propellant with the aid of a surfactant. The propellant may be any conventional material employed for this purpose, such as a chlorofluorocarbon, a hydrochlorofluorocarbon, a hydrofluorocarbon, or a hydrocarbon, including trichlorofluoromethane, dichlorodifluoromethane, dichlorotetrafluoroethanol, and 1,1,1,2-tetrafluoroethane, or combinations thereof. Suitable surfactants include sorbitan trioleate and soya lecithin. Oleic acid may also be useful as a surfactant.

Formulations for dispensing from a powder inhaler device will comprise a finely divided dry powder containing the inventive compound and may also include a bulking agent, such as lactose, sorbitol, sucrose, mannitol, trehalose, or xylitol in amounts which facilitate dispersal of the powder from the device, e.g., 50 to 90% by weight of the formulation.

Nasal delivery forms. Nasal delivery of the inventive compound is also contemplated. Nasal delivery allows the passage of the protein to the blood stream directly after administering the therapeutic product to the nose, without the necessity for deposition of the product in the lung. Formulations for nasal delivery include those with dextran or cyclodextran. Delivery via transport across other mucous membranes is also contemplated.

Dosages. The dosage regimen involved in a method for treating the above-described conditions will be determined by the attending physician, considering various factors which modify the action of drugs, e.g. the age, condition, body weight, sex and diet of the patient, the severity of any infection, time of administration and other clinical factors. Generally, the daily regimen should be in the range of 0.1-1000 micrograms of the inventive compound per kilogram of body weight, preferably 0.1-150 micrograms per kilogram.

Specific preferred embodiments

The inventors have determined preferred peptide sequences for molecules having many different kinds of activity. The inventors have further determined preferred structures of these preferred peptides combined with preferred linkers and vehicles. Preferred structures for these preferred peptides listed in Table 21 below.

Table 21—Preferred embodiments

Sequence/structure	SEQ ID NO:	Activity
F^1-$(G)_5$-IEGPTLRQWLAARA-$(G)_8$-IEGPTLRQWLAARA	337	TPO-mimetic
IEGPTLRQWLAARA-$(G)_8$-IEGPTLRQWLAARA-$(G)_5$- F^1	338	TPO-mimetic
F^1-$(G)_5$-IEGPTLRQWLAARA	1032	TPO-mimetic
IEGPTLRQWLAARA -$(G)_5$- F^1	1033	TPO-mimetic
F^1-$(G)_5$-GGTYSCHFGPLTWVCKPQGG-$(G)_4$-GGTYSCHFGPLTWVCKPQGG	339	EPO-mimetic
GGTYSCHFGPLTWVCKPQGG-$(G)_4$-GGTYSCHFGPLTWVCKPQGG-$(G)_5$-F^1	340	EPO-mimetic
GGTYSCHFGPLTWVCKPQGG-$(G)_5$-F^1	1034	EPO-mimetic
F^1-$(G)_5$-DFLPHYKNTSLGHRP	1045	TNF-α inhibitor
DFLPHYKNTSLGHRP-$(G)_5$-F^1	1046	TNF-α inhibitor
F^1-$(G)_5$- FEWTPGYWQPYALPL	1047	IL-1 R antagonist
FEWTPGYWQPYALPL-$(G)_5$-F^1	1048	IL-1 R antagonist
F^1-$(G)_5$-VEPNCDIHVMWEWECFERL	1049	VEGF-antagonist
VEPNCDIHVMWEWECFERL-$(G)_5$-F^1	1050	VEGF-antagonist
F^1-$(G)_5$-CTTHWGFTLC	1051	MMP inhibitor
CTTHWGFTLC-$(G)_5$-F^1	1052	MMP inhibitor

"F^1" is an Fc domain as defined previously herein.

Working examples

The compounds described above may be prepared as described below. These examples comprise preferred embodiments of the invention and are illustrative rather than limiting.

Example 1

TPO-Mimetics

The following example uses peptides identified by the numbers appearing in Table A hereinafter.

Preparation of peptide 19. Peptide 17b (12 mg) and MeO-PEG-SH 5000 (30 mg, 2 equiv.) were dissolved in 1 ml aqueous buffer (pH 8). The mixture was incubated at RT for about 30 minutes and the reaction was checked by analytical HPLC, which showed a > 80% completion of the reaction. The pegylated material was isolated by preparative HPLC.

Preparation of peptide 20. Peptide 18 (14 mg) and MeO-PEG-maleimide (25 mg) were dissolved in about 1.5 ml aqueous buffer (pH 8). The mixture was incubated at RT for about 30 minutes, at which time about 70% transformation was complete as monitored with analytical HPLC by applying an aliquot of sample to the HPLC column. The pegylated material was purified by preparative HPLC.

Bioactivity assay. The TPO in vitro bioassay is a mitogenic assay utilizing an IL-3 dependent clone of murine 32D cells that have been transfected with human mpl receptor. This assay is described in greater detail in WO 95/26746. Cells are maintained in MEM medium containing 10% Fetal Clone II and 1 ng/ml mIL-3. Prior to sample addition, cells are prepared by rinsing twice with growth medium lacking mIL-3. An extended twelve point TPO standard curve is prepared, ranging from 33 to 39 pg/ml. Four dilutions, estimated to fall within the linear portion of the standard curve, (100 to 125 pg/ml), are prepared for each sample and run in triplicate. A volume of 100 µl of each dilution of sample or standard is added to appropriate wells of a 96 well microtiter plate

containing 10,000 cells/well. After forty-four hours at 37 °C and 10% CO_2, MTS (a tetrazolium compound which is bioreduced by cells to a formazan) is added to each well. Approximately six hours later, the optical density is read on a plate reader at 490 nm. A dose response curve (log TPO concentration vs. O.D.- Background) is generated and linear regression analysis of points which fall in the linear portion of the standard curve is performed. Concentrations of unknown test samples are determined using the resulting linear equation and a correction for the dilution factor.

TMP tandem repeats with polyglycine linkers. Our design of sequentially linked TMP repeats was based on the assumption that a dimeric form of TMP was required for its effective interaction with c-Mpl (the TPO receptor) and that depending on how they were wound up against each other in the receptor context, the two TMP molecules could be tethered together in the C- to N-terminus configuration in a way that would not perturb the global dimeric conformation. Clearly, the success of the design of tandem linked repeats depends on proper selection of the length and composition of the linker that joins the C- and N-termini of the two sequentially aligned TMP monomers. Since no structural information of the TMP bound to c-Mpl was available, a series of repeated peptides with linkers composed of 0 to 10 and 14 glycine residues (Table A) were synthesized. Glycine was chosen because of its simplicity and flexibility, based on the rationale that a flexible polyglycine peptide chain might allow for the free folding of the two tethered TMP repeats into the required conformation, while other amino acid sequences may adopt undesired secondary structures whose rigidity might disrupt the correct packing of the repeated peptide in the receptor context.

The resulting peptides are readily accessible by conventional solid phase peptide synthesis methods (Merrifield (1963), J. Amer. Chem. Soc. 85: 2149) with either Fmoc or t-Boc chemistry. Unlike the synthesis of the

C-terminally linked parallel dimer which required the use of an orthogonally protected lysine residue as the initial branch point to build the two peptide chains in a pseudosymmetrical way (Cwirla et al. (1997), Science 276: 1696-9), the synthesis of these tandem repeats was a straightforward, stepwise assembly of the continuous peptide chains from the C- to N-terminus. Since dimerization of TMP had a more dramatic effect on the proliferative activity than binding affinity as shown for the C-terminal dimer (Cwirla et al. (1997)), the synthetic peptides were tested directly for biological activity in a TPO-dependent cell-proliferation assay using an IL-3 dependent clone of murine 32D cells transfected with the full-length c-Mpl (Palacios et al.,. Cell 41:727 (1985)). As the test results showed, all the polyglycine linked tandem repeats demonstrated >1000 fold increases in potency as compared to the monomer, and were even more potent than the C-terminal dimer in this cell proliferation assay. The absolute activity of the C-terminal dimer in our assay was lower than that of the native TPO protein, which is different from the previously reported findings in which the C-terminal dimer was found to be as active as the natural ligand (Cwirla et al. (1997)). This might be due to differences in the conditions used in the two assays. Nevertheless, the difference in activity between tandem (C terminal of first monomer linked to N terminal of second monomer) and C-terminal (C terminal of first monomer linked to C terminal of second monomer; also referred to as parallel) dimers in the same assay clearly demonstrated the superiority of tandem repeat strategy over parallel peptide dimerization. It is interesting to note that a wide range of length is tolerated by the linker. The optimal linker between tandem peptides with the selected TMP monomers apparently is composed of 8 glycines.

Other tandem repeats. Subsequent to this first series of TMP tandem repeats, several other molecules were designed either with

different linkers or containing modifications within the monomer itself. The first of these molecules, peptide 13, has a linker composed of GPNG, a sequence known to have a high propensity to form a β-turn-type secondary structure. Although still about 100-fold more potent than the monomer, this peptide was found to be >10-fold less active than the equivalent GGGG-linked analog. Thus, introduction of a relatively rigid β-turn at the linker region seemed to have caused a slight distortion of the optimal agonist conformation in this short linker form.

The Trp9 in the TMP sequence is a highly conserved residue among the active peptides isolated from random peptide libraries. There is also a highly conserved Trp in the consensus sequences of EPO mimetic peptides and this Trp residue was found to be involved in the formation of a hydrophobic core between the two EMPs and contributed to hydrophobic interactions with the EPO receptor. Livnah et al. (1996), Science 273: 464-71). By analogy, the Trp9 residue in TMP might have a similar function in dimerization of the peptide ligand, and as an attempt to modulate and estimate the effects of noncovalent hydrophobic forces exerted by the two indole rings, several analogs were made resulting from mutations at the Trp. So in peptide 14, the Trp residue was replaced in each of the two TMP monomers with a Cys, and an intramolecular disulfide bond was formed between the two cysteines by oxidation which was envisioned to mimic the hydrophobic interactions between the two Trp residues in peptide dimerization. Peptide 15 is the reduced form of peptide 14. In peptide 16, the two Trp residues were replaced by Ala. As the assay data show, all three analogs were inactive. These data further demonstrated that Trp is critical for the activity of the TPO mimetic peptide, not just for dimer formation.

The next two peptides (peptide 17a, and 18) each contain in their 8-amino acid linker a Lys or Cys residue. These two compounds are

precursors to the two PEGylated peptides (peptide 19 and 20) in which the side chain of the Lys or Cys is modified by a PEG moiety. A PEG moiety was introduced at the middle of a relatively long linker, so that the large PEG component (5 kDa) is far enough away from the critical binding sites in the peptide molecule. PEG is a known biocompatible polymer which is increasingly used as a covalent modifier to improve the pharmacokinetic profiles of peptide- and protein-based therapeutics.

A modular, solution-based method was devised for convenient PEGylation of synthetic or recombinant peptides. The method is based on the now well established chemoselective ligation strategy which utilizes the specific reaction between a pair of mutually reactive functionalities. So, for pegylated peptide 19, the lysine side chain was preactivated with a bromoacetyl group to give peptide 17b to accommodate reaction with a thiol-derivatized PEG. To do that, an orthogonal protecting group, Dde, was employed for the protection of the lysine ε-amine. Once the whole peptide chain was assembled, the N-terminal amine was reprotected with t-Boc. Dde was then removed to allow for the bromoacetylation. This strategy gave a high quality crude peptide which was easily purified using conventional reverse phase HPLC. Ligation of the peptide with the thiol-modified PEG took place in aqueous buffer at pH 8 and the reaction completed within 30 minutes. MALDI-MS analysis of the purified, pegylated material revealed a characteristic, bell-shaped spectrum with an increment of 44 Da between the adjacent peaks. For PEG-peptide 20, a cysteine residue was placed in the linker region and its side chain thiol group would serve as an attachment site for a maleimide-containing PEG. Similar conditions were used for the pegylation of this peptide. As the assay data revealed, these two pegylated peptides had even higher in vitro bioactivity as compared to their unpegylated counterparts.

Peptide 21 has in its 8-amino acid linker a potential glycosylation motif, NGS. Since our exemplary tandem repeats are made up of natural amino acids linked by peptide bonds, expression of such a molecule in an appropriate eukaryotic cell system should produce a glycopeptide with the carbohydrate moiety added on the side chain carboxyamide of Asn. Glycosylation is a common post-translational modification process which can have many positive impacts on the biological activity of a given protein by increasing its aqueous solubility and in vivo stability. As the assay data show, incorporation of this glycosylation motif into the linker maintained high bioactivity. The synthetic precursor of the potential glycopeptide had in effect an activity comparable to that of the -$(G)_8$- linked analog. Once glycosylated, this peptide is expected to have the same order of activity as the pegylated peptides, because of the similar chemophysical properties exhibited by a PEG and a carbohydrate moiety.

The last peptide is a dimer of a tandem repeat. It was prepared by oxidizing peptide 18, which formed an intermolecular disulfide bond between the two cysteine residues located at the linker. This peptide was designed to address the possibility that TMP was active as a tetramer. The assay data showed that this peptide was not more active than an average tandem repeat on an adjusted molar basis, which indirectly supports the idea that the active form of TMP is indeed a dimer, otherwise dimerization of a tandem repeat would have a further impact on the bioactivity.

In order to confirm the in vitro data in animals, one pegylated TMP tandem repeat (compound 20 in Table A) was delivered subcutaneously to normal mice via osmotic pumps. Time and dose-dependent increases were seen in platelet numbers for the duration of treatment. Peak platelet levels over 4-fold baseline were seen on day 8. A dose of 10 μg/kg/day of the pegylated TMP repeat produced a similar response to rHuMGDF (non-pegylated) at 100 μg/kg/day delivered by the same route.

Table A—TPO-mimetic Peptides

Peptide No.	Compound	SEQ ID NO:	Relative Potency
	TPO		++++
	TMP monomer	13	+
	TMP C-C dimer		+++-
TMP-$(G)_n$-TMP:			
1	n = 0	341	++++-
2	n = 1	342	++++
3	n = 2	343	++++
4	n = 3	344	++++
5	n = 4	345	++++
6	n = 5	346	++++
7	n = 6	347	++++
8	n = 7	348	++++
9	n = 8	349	++++-
10	n = 9	350	++++
11	n = 10	351	++++
12	n = 14	352	++++
13	TMP-GPNG-TMP	353	+++
14	IEGPTLRQCLAARA-GGGGGGGG-IEGPTLRQCLAARA (cyclic)	354	-
15	IEGPTLRQCLAARA-GGGGGGGG-IEGPTLRQCLAARA (linear)	355	-
16	IEGPTLRQALAARA-GGGGGGGG-IEGPTLRQALAARA	356	-
17a	TMP-GGGKGGGG-TMP	357	++++
17b	TMP-GGGK(BrAc)GGGG-TMP	358	ND
18	TMP-GGGCGGGG-TMP	359	++++
19	TMP-GGGK(PEG)GGGG-TMP	360	+++++
20	TMP-GGGC(PEG)GGGG-TMP	361	+++++
21	TMP-GGGN*GSGG-TMP	362	++++
22	TMP-GGGCGGGG-TMP | TMP-GGGCGGGG-TMP	363 363	++++

Discussion. It is well accepted that MGDF acts in a way similar to hGH, i.e., one molecule of the protein ligand binds two molecules of the receptor for its activation. Wells et al.(1996), Ann. Rev. Biochem. 65: 609-34. Now, this interaction is mimicked by the action of a much smaller peptide, TMP. However, the present studies suggest that this mimicry requires the concerted action of two TMP molecules, as covalent dimerization of TMP in either a C-C parallel or C-N sequential fashion increased the in vitro biological potency of the original monomer by a factor of greater than 10^3. The relatively low biopotency of the monomer is probably due to inefficient formation of the noncovalent dimer. A preformed covalent repeat has the ability to eliminate the entropy barrier for the formation of a noncovalent dimer which is exclusively driven by weak, noncovalent interactions between two molecules of the small, 14-residue peptide.

It is intriguing that this tandem repeat approach had a similar effect on enhancing bioactivity as the reported C-C dimerization is intriguing. These two strategies brought about two very different molecular configurations. The C-C dimer is a quasi-symmetrical molecule, while the tandem repeats have no such symmetry in their linear structures. Despite this difference in their primary structures, these two types of molecules appeared able to fold effectively into a similar biologically active conformation and cause the dimerization and activation of c-Mpl. These experimental observations provide a number of insights into how the two TMP molecules may interact with one another in binding to c-Mpl. First, the two C-termini of the two bound TMP molecules must be in relatively close proximity with each other, as suggested by data on the C-terminal dimer. Second, the respective N- and C-termini of the two TMP molecules in the receptor complex must also be very closely aligned with each other, such that they can be directly tethered together with a single peptide bond

to realize the near maximum activity-enhancing effect brought about by the tandem repeat strategy. Insertion of one or more (up to 14) glycine residues at the junction did not increase (or decrease) significantly the activity any further. This may be due to the fact that a flexible polyglycine peptide chain is able to loop out easily from the junction without causing any significant changes in the overall conformation. This flexibility seems to provide the freedom of orientation for the TMP peptide chains to fold into the required conformation in interacting with the receptor and validate it as a site of modification. Indirect evidence supporting this came from the study on peptide 13, in which a much more rigid b-turn-forming sequence as the linker apparently forced a deviation of the backbone alignment around the linker which might have resulted in a slight distortion of the optimal conformation, thus resulting in a moderate (10-fold) decrease in activity as compared with the analogous compound with a 4-Gly linker. Third, Trp9 in TMP plays a similar role as Trp13 in EMP, which is involved not only in peptide:peptide interaction for the formation of dimers but also is important for contributing hydrophobic forces in peptide:receptor interaction. Results obtained with the W to C mutant analog, peptide 14, suggest that a covalent disulfide linkage is not sufficient to approximate the hydrophobic interactions provided by the Trp pair and that, being a short linkage, it might bring the two TMP monomers too close, therefore perturbing the overall conformation of the optimal dimeric structure.

An analysis of the possible secondary structure of the TMP peptide can provide further understanding on the interaction between TMP and c-Mpl. This can be facilitated by making reference to the reported structure of the EPO mimetic peptide. Livnah et al. (1996), Science 273:464-75 The receptor-bound EMP has a b-hairpin structure with a b-turn formed by the highly consensus Gly-Pro-Leu-Thr at the center of its sequence. Instead of

GPLT, TMP has a highly selected GPTL sequence which is likely to form a similar turn. However, this turn-like motif is located near the N-terminal part in TMP. Secondary structure prediction using Chau-Fasman method suggests that the C-terminal half of the peptide has a tendency to adopt a helical conformation. Together with the highly conserved Trp at position 9, this C-terminal helix may contribute to the stabilization of the dimeric structure. It is interesting to note that most of our tandem repeats are more potent than the C-terminal parallel dimer. Tandem repeats seem to give the molecule a better fit conformation than does the C-C parallel dimerization. The seemingly asymmetric feature of a tandem repeat might have brought it closer to the natural ligand which, as an asymmetric molecule, uses two different sites to bind two identical receptor molecules.

Introduction of a PEG moiety was envisaged to enhance the in vivo activity of the modified peptide by providing it a protection against proteolytic degradation and by slowing down its clearance through renal filtration. It was unexpected that pegylation could further increase the in vitro bioactivity of a tandem repeated TMP peptide in the cell-based proliferation assay.

Example 2

Fc-TMP fusions

TMPs (and EMPs as described in Example 3) were expressed in either monomeric or dimeric form as either N-terminal or C-terminal fusions to the Fc region of human IgG1. In all cases, the expression construct utilized the luxPR promoter promoter in the plasmid expression vector pAMG21.

Fc-TMP. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of the TPO-mimetic peptide was constructed using standard PCR technology. Templates for PCR reactions were the pFc-A3 vector and a synthetic TMP gene. The synthetic gene was

constructed from the 3 overlapping oligonucleotides (SEQ ID NOS: 364, 365, and 366, respectively) shown below:

```
1842-97      AAA AAA GGA TCC TCG AGA TTA AGC ACG AGC AGC CAG CCA
             CTG ACG CAG AGT CGG ACC

1842-98      AAA GGT GGA GGT GGT GGT ATC GAA GGT CCG ACT CTG CGT

1842-99      CAG TGG CTG GCT GCT CGT GCT TAA TCT CGA GGA TCC TTT
             TTT
```

These oligonucleotides were annealed to form the duplex encoding an amino acid sequence (SEQ ID NOS: 367 and 368, respectively) shown below:

```
          AAAGGTGGAGGTGGTGGTATCGAAGGTCCGACTCTGCGTCAGTGGCTGGCTGCTCGTGCT
        1 ---------+---------+---------+---------+---------+---------+ 60
                                        CCAGGCTGAGACGCAGTCACCGACCGACGAGCACGA
a         K  G  G  G  G  G  I  E  G  P  T  L  R  Q  W  L  A  A  R  A   -

          TAATCTCGAGGATCCTTTTTT
       61 ---------+---------+- 81
          ATTAGAGCTCCTAGGAAAAAA
a         *
```

This duplex was amplified in a PCR reaction using 1842-98 and 1842-97 as the sense and antisense primers.

The Fc portion of the molecule was generated in a PCR reaction with pFc-A3 using the primers shown below (SEQ ID NOS: 369 and 370):

```
1216-52      AAC ATA AGT ACC TGT AGG ATC G

1830-51      TTCGATACCA CCACCTCCAC CTTTACCCGG AGACAGGGAG AGGCTCTTCTGC
```

The oligonucleotides 1830-51 and 1842-98 contain an overlap of 24 nucleotides, allowing the two genes to be fused together in the correct reading frame by combining the above PCR products in a third reaction using the outside primers, 1216-52 and 1842-97.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the

gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3728.

The nucleotide and amino acid sequences (SEQ ID NOS: 5 and 6) of the fusion protein are shown in Figure 7.

Fc-TMP-TMP. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a dimer of the TPO-mimetic peptide was constructed using standard PCR technology. Templates for PCR reactions were the pFc-A3 vector and a synthetic TMP-TMP gene. The synthetic gene was constructed from the 4 overlapping oligonucleotides (SEQ ID NOS: 371 to 374, respectively) shown below:

```
1830-52     AAA GGT GGA GGT GGT GGT ATC GAA GGT CCG
            ACT CTG CGT CAG TGG CTG GCT GCT CGT GCT

1830-53     ACC TCC ACC ACC AGC ACG AGC AGC CAG
            CCA CTG ACG CAG AGT CGG ACC

1830-54     GGT GGT GGA GGT GGC GGC GGA GGT ATT GAG GGC CCA ACC
            CTT CGC CAA TGG CTT GCA GCA CGC GCA

1830-55     AAA AAA AGG ATC CTC GAG ATT ATG CGC GTG CTG CAA GCC
            ATT GGC GAA GGG TTG GGC CCT CAA TAC CTC CGC CGC C
```

The 4 oligonucleotides were annealed to form the duplex encoding an amino acid sequence (SEQ ID NOS: 375 and 376, respectively) shown below:

```
          AAAGGTGGAGGTGGTGGTATCGAAGGTCCGACTCTGCGTCAGTGGCTGGCTGCTCGTGCT
        1 ---------+---------+---------+---------+---------+---------+ 60
                                    CCAGGCTGAGACGCAGTCACCGACCGACGAGCACGA
a         K  G  G  G  G  G  I  E  G  P  T  L  R  Q  W  L  A  A  R  A   -

          GGTGGTGGAGGTGGCGGCGGAGGTATTGAGGGCCCAACCCTTCGCCAATGGCTTGCAGCA
       61 ---------+---------+---------+---------+---------+---------+ 120
          CCACCACCTCCACCGCCGCCTCCATAACTCCCGGGTTGGGAAGCGGTTACCGAACGTCGT

a         G  G  G  G  G  G  G  I  E  G  P  T  L  R  Q  W  L  A  A   -
          CGCGCA
      121 ------------------------148
          GCGCGTATTAGAGCTCCTAGGAAAAAAA

a         R  A  *-
```

This duplex was amplified in a PCR reaction using 1830-52 and 1830-55 as the sense and antisense primers.

The Fc portion of the molecule was generated in a PCR reaction with pFc-A3 using the primers 1216-52 and 1830-51 as described above for

Fc-TMP. The full length fusion gene was obtained from a third PCR reaction using the outside primers 1216-52 and 1830-55.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described in example 1. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3727.

The nucleotide and amino acid sequences (SEQ ID NOS: 7 and 8) of the fusion protein are shown in Figure 8.

TMP-TMP-Fc. A DNA sequence coding for a tandem repeat of the TPO-mimetic peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. Templates for PCR reactions were the EMP-Fc plasmid from strain #3688 (see Example 3) and a synthetic gene encoding the TMP dimer. The synthetic gene for the tandem repeat was constructed from the 7 overlapping oligonucleotides shown below (SEQ ID NOS: 377 to 383, respectively):

```
1885-52    TTT TTT CAT ATG ATC GAA GGT CCG ACT CTG CGT CAG TGG

1885-53    AGC ACG AGC AGC CAG CCA CTG ACG CAG AGT CGG ACC TTC
           GAT CAT ATG

1885-54    CTG GCT GCT CGT GCT GGT GGA GGC GGT GGG GAC AAA ACT
           CAC ACA

1885-55    CTG GCT GCT CGT GCT GGC GGT GGT GGC GGA GGG GGT GGC
           ATT GAG GGC CCA

1885-56    AAG CCA TTG GCG AAG GGT TGG GCC CTC AAT GCC ACC CCC
           TCC GCC ACC ACC GCC

1885-57    ACC CTT CGC CAA TGG CTT GCA GCA CGC GCA GGG GGA GGC
           GGT GGG GAC AAA ACT

1885-58    CCC ACC GCC TCC CCC TGC GCG TGC TGC
```

These oligonucleotides were annealed to form the duplex shown encoding an amino acid sequence shown below (SEQ ID NOS 384 and 385):

```
          TTTTTTCATATGATCGAAGGTCCGACTCTGCGTCAGTGGCTGGCTGCTCGTGCTGGCGGT
        1 ---------+---------+---------+---------+---------+---------+ 60
                GTATACTAGCTTCCAGGCTGAGACGCAGTCACCGACCGACGAGCACGACCGCCA
a                   M  I  E  G  P  T  L  R  Q  W  L  A  A  R  A  G  G   -

          GGTGGCGGAGGGGGTGGCATTGAGGGCCCAACCCTTCGCCAATGGCTGGCTGCTCGTGCT
       61 ---------+---------+---------+---------+---------+---------+ 120
          CCACCGCCTCCCCCACCGTAACTCCCGGGTTGGGAAGCGGTTACCGAACGTCGTGCGCGT
a         G  G  G  G  G  G  I  E  G  P  T  L  R  Q  W  L  A  A  R  A   -

          GGTGGAGGCGGTGGGGACAAAACTCTGGCTGCTCGTGCTGGTGGAGGCGGTGGGGACAAA
      121 ---------+---------+---------+---------+---------+---------+ 180
          CCCCCTCCGCCACCC
a         G  G  G  G  G  D  K  T  L  A  A  R  A  G  G  G  G  D  K   -

          ACTCACACA
      181 --------- 189

a         T  H  T   -
```

This duplex was amplified in a PCR reaction using 1885-52 and 1885-58 as the sense and antisense primers.

The Fc portion of the molecule was generated in a PCR reaction with DNA from the EMP-Fc fusion strain #3688 (see Example 3) using the primers 1885-54 and 1200-54. The full length fusion gene was obtained from a third PCR reaction using the outside primers 1885-52 and 1200-54.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for Fc-EMP herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3798.

The nucelotide and amino acid sequences (SEQ ID NOS: 9 and 10) of the fusion protein are shown in Figure 9.

TMP-Fc. A DNA sequence coding for a monomer of the TPO-mimetic peptide fused in-frame to the Fc region of human IgG1 was obtained fortuitously in the ligation in TMP-TMP-Fc, presumably due to the ability of primer 1885-54 to anneal to 1885-53 as well as to 1885-58. A single clone having the correct nucleotide sequence for the TMP-Fc construct was selected and designated Amgen strain #3788.

The nucleotide and amino acid sequences (SEQ ID NOS: 11 and 12) of the fusion protein are shown in Figure 10.

Expression in E. coli. Cultures of each of the pAMG21-Fc-fusion constructs in E. coli GM221 were grown at 37 °C in Luria Broth medium containing 50 mg/ml kanamycin. Induction of gene product expression from the luxPR promoter was achieved following the addition of the synthetic autoinducer N-(3-oxohexanoyl)-DL-homoserine lactone to the culture media to a final concentration of 20 ng/ml. Cultures were incubated at 37 °C for a further 3 hours. After 3 hours, the bacterial cultures were examined by microscopy for the presence of inclusion bodies and were then collected by centrifugation. Refractile inclusion bodies were observed in induced cultures indicating that the Fc-fusions were most likely produced in the insoluble fraction in E. coli. Cell pellets were lysed directly by resuspension in Laemmli sample buffer containing 10% b-mercaptoethanol and were analyzed by SDS-PAGE. In each case, an intense coomassie-stained band of the appropriate molecular weight was observed on an SDS-PAGE gel.

pAMG21. The expression plasmid pAMG21 can be derived from the Amgen expression vector pCFM1656 (ATCC #69576) which in turn be derived from the Amgen expression vector system described in US Patent No. 4,710,473. The pCFM1656 plasmid can be derived from the described pCFM836 plasmid (Patent No. 4,710,473) by:

(a) destroying the two endogenous NdeI restriction sites by end filling with T4 polymerase enzyme followed by blunt end ligation;

(b) replacing the DNA sequence between the unique AatII and ClaI restriction sites containing the synthetic P_L promoter with a similar fragment obtained from pCFM636 (patent No. 4,710,473) containing the PL promoter (see SEQ ID NO: 386 below); and

(c) substituting the small DNA sequence between the unique ClaI and KpnI restriction sites with the oligonucleotide having the sequence of SEQ ID NO: 388.

SEQ ID NO: 386:

```
     AatII
5'  CTAATTCCGCTCTCACCTACCAAACAATGCCCCCCTGCAAAAAATAAATTCATAT-
3'  TGCAGATTAAGGCGAGAGTGGATGGTTTGTTACGGGGGGACGTTTTTTATTTAAGTATA-

    -AAAAAACATACAGATAACCATCTGCGGTGATAAATTATCTCTGGCGGTGTTGACATAAA-
    -TTTTTTGTATGTCTATTGGTAGACGCCACTATTTAATAGAGACCGCCACAACTGTATTT-

    -TACCACTGGCGGTGATACTGAGCACAT    3'
    -ATGGTGACCGCCACTATGACTCGTGTAGC  5'
                                ClaI
```

SEQ ID NO: 387:

```
5'  CGATTTGATTCTAGAAGGAGGAATAACATATGGTTAACGCGTTGGAATTCGGTAC  3'
3'    TAAACTAAGATCTTCCTCCTTATTGTATACCAATTGCGCAACCTTAAGC       5'
   ClaI                                                  KpnI
```

The expression plasmid pAMG21 can then be derived from pCFM1656 by making a series of site-directed base changes by PCR overlapping oligo mutagenesis and DNA sequence substitutions. Starting with the BglII site (plasmid bp # 180) immediately 5' to the plasmid replication promoter P_{copB} and proceeding toward the plasmid replication genes, the base pair changes are as shown in Table B below.

Table B—Base pair changes resulting in pAMG21

pAMG21 bp #	bp in pCFM1656	bp changed to in pAMG21
# 204	T/A	C/G
# 428	A/T	G/C
# 509	G/C	A/T
# 617	- -	insert two G/C bp
# 679	G/C	T/A
# 980	T/A	C/G
# 994	G/C	A/T
# 1004	A/T	C/G
# 1007	C/G	T/A
# 1028	A/T	T/A
# 1047	C/G	T/A
# 1178	G/C	T/A
# 1466	G/C	T/A
# 2028	G/C	bp deletion
# 2187	C/G	T/A
# 2480	A/T	T/A
# 2499-2502	AGTG TCAC	GTCA CAGT
# 2642	TCCGAGC AGGCTCG	7 bp deletion
# 3435	G/C	A/T
# 3446	G/C	A/T
# 3643	A/T	T/A

The DNA sequence between the unique AatII (position #4364 in pCFM1656) and SacII (position #4585 in pCFM1656) restriction sites is substituted with the DNA sequence (SEQ ID NO: 23) shown in Figures 17A and 17B. During the ligation of the sticky ends of this substitution DNA sequence, the outside AatII and SacII sites are destroyed. There are unique AatII and SacII sites in the substituted DNA.

GM221 (Amgen #2596). The Amgen host strain #2596 is an E.coli K-12 strain derived from Amgen strain #393. It has been modified to contain both the temperature sensitive lambda repressor cI857s7 in the early ebg region and the $lacI^Q$ repressor in the late ebg region (68 minutes). The presence of these two repressor genes allows the use of this host with a variety of expression systems, however both of these repressors are irrelevant to the expression from $luxP_R$. The untransformed host has no antibiotic resistances.

The ribosome binding site of the cI857s7 gene has been modified to include an enhanced RBS. It has been inserted into the ebg operon between nucleotide position 1170 and 1411 as numbered in Genbank accession number M64441Gb_Ba with deletion of the intervening ebg sequence. The sequence of the insert is shown below with lower case letters representing the ebg sequences flanking the insert shown below (SEQ ID NO: 388):

```
ttatttcgtGCGGCCGCACCATTATCACCGCCAGAGGTAAACTAGTCAACACGCACGGTGTTAGATATTTAT
CCCTTGCGGTGATAGATTGAGCACATCGATTTGATTCTAGAAGGAGGGATAATATATGAGCACAAAAAAGAAA
CCATTAACACAAGAGCAGCTTGAGGACGCACGTCGCCTTAAAGCAATTTATGAAAAAAAGAAAAATGAACTTG
GCTTATCCCAGGAATCTGTCGCAGACAAGATGGGGATGGGGCAGTCAGGCGTTGGTGCTTTATTTAATGGCAT
CAATGCATTAAATGCTTATAACGCCGCATTGCTTACAAAAATTCTCAAAGTTAGCGTTGAAGAATTTAGCCCT
TCAATCGCCAGAGAATCTACGAGATGTATGAAGCGGTTAGTATGCAGCCGTCACTTAGAAGTGAGTATGAGTA
CCCTGTTTTTTCTCATGTTCAGGCAGGGATGTTCTCACCTAAGCTTAGAACCTTTACCAAAGGTGATGCGGAG
AGATGGGTAAGCACAACCAAAAAAGCCAGTGATTCTGCATTCTGGCTTGAGGTTGAAGGTAATTCCATGACCG
CACCAACAGGCTCCAAGCCAAGCTTTCCTGACGGAATGTTAATTCTCGTTGACCCTGAGCAGGCTGTTGAGCC
AGGTGATTTCTGCATAGCCAGACTTGGGGGTGATGAGTTTACCTTCAAGAAACTGATCAGGGATAGCGGTCAG
GTGTTTTTACAACCACTAAACCCACAGTACCCAATGATCCCATGCAATGAGAGTTGTTCCGTTGTGGGGAAAG
TTATCGCTAGTCAGTGGCCTGAAGAGACGTTTGGCTGATAGACTAGTGGATCCACTAGTgtttctgccc
```

The construct was delivered to the chromosome using a recombinant phage called MMebg-cI857s7enhanced RBS #4 into F'tet/393. After recombination and resolution only the chromosomal insert described

above remains in the cell. It was renamed F'tet/GM101. F'tet/GM101 was then modified by the delivery of a $lacI^Q$ construct into the ebg operon between nucleotide position 2493 and 2937 as numbered in the Genbank accession number M64441Gb_Ba with the deletion of the intervening ebg sequence. The sequence of the insert is shown below with the lower case letters representing the ebg sequences flanking the insert (SEQ ID NO: 389) shown below:

```
ggcggaaaccGACGTCCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA
ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACC
GTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTCGAAGCGGCGATGGCGG
AGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGCTCCTGATTGGCGTTGCCAC
CTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCC
AGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGC
AACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCAC
TAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGAC
GGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAA
GTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGC
GGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTT
CCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGC
GCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAC
CACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAG
GCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAA
CCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGACA
GTAAGGTACCATAGGATCCaggcacagga
```

The construct was delivered to the chromosome using a recombinant phage called AGebg-LacIQ#5 into F'tet/GM101. After recombination and resolution only the chromosomal insert described above remains in the cell. It was renamed F'tet/GM221. The F'tet episome was cured from the strain using acridine orange at a concentration of 25 μg/ml in LB. The cured strain was identified as tetracyline sensitive and was stored as GM221.

Expression. Cultures of pAMG21-Fc-TMP-TMP in *E. coli* GM221 in Luria Broth medium containing 50 μg/ml kanamycin were incubated at 37°C prior to induction. Induction of Fc-TMP-TMP gene product expression from the luxPR promoter was achieved following the addition of the synthetic autoinducer N-(3-oxohexanoyl)-DL-homoserine lactone to the culture media to a final concentration of 20 ng/ml and cultures were incubated at 37°C for a further 3 hours. After 3 hours, the bacterial

cultures were examined by microscopy for the presence of inclusion bodies and were then collected by centrifugation. Refractile inclusion bodies were observed in induced cultures indicating that the Fc-TMP-TMP was most likely produced in the insoluble fraction in *E. coli*. Cell pellets were lysed directly by resuspension in Laemmli sample buffer containing 10% •-mercaptoethanol and were analyzed by SDS-PAGE. An intense Coomassie stained band of approximately 30kDa was observed on an SDS-PAGE gel. The expected gene product would be 269 amino acids in length and have an expected molecular weight of about 29.5 kDa. Fermentation was also carried out under standard batch conditions at the 10 L scale, resulting in similar expression levels of the Fc-TMP-TMP to those obtained at bench scale.

Purification of Fc-TMP-TMP. Cells are broken in water (1/10) by high pressure homogenization (2 passes at 14,000 PSI) and inclusion bodies are harvested by centrifugation (4200 RPM in J-6B for 1 hour). Inclusion bodies are solubilized in 6M guanidine, 50mM Tris, 8mM DTT, pH 8.7 for 1 hour at a 1/10 ratio. The solubilized mixture is diluted 20 times into 2M urea, 50 mM tris, 160mM arginine, 3mM cysteine, pH 8.5. The mixture is stirred overnight in the cold and then concentrated about 10 fold by ultafiltration. It is then diluted 3 fold with 10mM Tris, 1.5M urea, pH 9. The pH of this mixture is then adjusted to pH 5 with acetic acid. The precipitate is removed by centrifugation and the supernatant is loaded onto a SP-Sepharose Fast Flow column equilibrated in 20mM NaAc, 100 mM NaCl, pH 5(10mg/ml protein load, room temperature). The protein is eluted off using a 20 column volume gradient in the same buffer ranging from 100mM NaCl to 500mM NaCl. The pool from the column is diluted 3 fold and loaded onto a SP-Sepharose HP column in 20 mM NaAc, 150 mM NaCl, pH 5(10 mg/ml protein load, room temperature). The protein is eluted off using a 20 column volume gradient

in the same buffer ranging from 150 mM NaCl to 400 mM NaCl. The peak is pooled and filtered.

Characterization of Fc-TMP activity. The following is a summary of in vivo data in mice with various compounds of this invention.

Mice: Normal female BDF1 approximately 10-12 weeks of age.

Bleed schedule: Ten mice per group treated on day 0, two groups started 4 days apart for a total of 20 mice per group. Five mice bled at each time point, mice were bled a minimum of three times a week. Mice were anesthetized with isoflurane and a total volume of 140-160 µl of blood was obtained by puncture of the orbital sinus. Blood was counted on a Technicon H1E blood analyzer running software for murine blood. Parameters measured were white blood cells, red blood cells, hematocrit, hemoglobin, platelets, neutrophils.

Treatments: Mice were either injected subcutaneously for a bolus treatment or implanted with 7-day micro-osmotic pumps for continuous delivery. Subcutaneous injections were delivered in a volume of 0.2 ml. Osmotic pumps were inserted into a subcutaneous incision made in the skin between the scapulae of anesthetized mice. Compounds were diluted in PBS with 0.1% BSA. All experiments included one control group, labeled "carrier" that were treated with this diluent only. The concentration of the test articles in the pumps was adjusted so that the calibrated flow rate from the pumps gave the treatment levels indicated in the graphs.

Compounds: A dose titration of the compound was delivered to mice in 7 day micro-osmotic pumps. Mice were treated with various compounds at a single dose of 100 µg/kg in 7 day osmotic pumps. Some of the same compounds were then given to mice as a single bolus injection.

Activity test results: The results of the activity experiments are shown in Figures 11 and 12. In dose response assays using 7-day micro-

osmotic pumps, the maximum effect was seen with the compound of SEQ ID NO: 18 was at 100 μg/kg/day; the 10 μg/kg/day dose was about 50% maximally active and 1 μg/kg/day was the lowest dose at which activity could be seen in this assay system. The compound at 10 μg/kg/day dose was about equally active as 100 μg/kg/day unpegylated rHu-MGDF in the same experiment.

Example 3

Fc-EMP fusions

Fc-EMP. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of the EPO-mimetic peptide was constructed using standard PCR technology. Templates for PCR reactions were a vector containing the Fc sequence (pFc-A3, described in International application WO 97/23614, published July 3, 1997) and a synthetic gene encoding EPO monomer. The synthetic gene for the monomer was constructed from the 4 overlapping oligonucleotides (SEQ ID NOS: 390 to 393, respectively) shown below:

```
1798-2 TAT GAA AGG TGG AGG TGG TGG TGG AGG TAC TTA CTC TTG
       CCA CTT CGG CCC GCT GAC TTG G

1798-3 CGG TTT GCA AAC CCA AGT CAG CGG GCC GAA GTG GCA AGA
       GTA AGT ACC TCC ACC ACC ACC TCC ACC TTT CAT

1798-4 GTT TGC AAA CCG CAG GGT GGC GGC GGC GGC GGT GGT
       ACC TAT TCC TGT CAT TTT

1798-5 CCA GGT CAG CGG GCC AAA ATG ACA GGA ATA GGT ACC ACC
       GCC GCC GCC GCC GCC ACC CTG
```

The 4 oligonucleotides were annealed to form the duplex encoding an amino acid sequence (SEQ ID NOS: 394 and 395, respectively) shown below:

```
   TATGAAAGGTGGAGGTGGTGGTGGAGGTACTTACTCTTGCCACTTCGGCCCGCTGACTTG
 1 ---------+---------+---------+---------+---------+---------+ 60
   TACTTTCCACCTCCACCACCACCTCCATGAATGAGAACGGTGAAGCCGGGCGACTGAAC
b   M  K  G  G  G  G  G  G  G  T  Y  S  C  H  F  G  P  L  T  W  -

   GGTTTGCAAACCGCAGGGTGGCGGCGGCGGCGGTGGTACCTATTCCTGTCATTTT
61 ---------+---------+---------+---------+---------+---------+-- 133
   CCAAACGTTTGGCGTCCCACCGCCGCCGCCGCCGCCACCATGGATAAGGACAGTAAAACCGGGCGACTGGACC
b   V  C  K  P  Q  G  G  G  G  G  G  G  G  T  Y  S  C  H  F   -
```

This duplex was amplified in a PCR reaction using

```
1798-18      GCA GAA GAG CCT CTC CCT GTC TCC GGG TAA
             AGG TGG AGG TGG TGG TGG AGG TAC TTA
             CTC T
```

and

```
1798-19      CTA ATT GGA TCC ACG AGA TTA ACC ACC
             CTG CGG TTT GCA A
```

as the sense and antisense primers (SEQ ID NOS: 396 and 397, respectively).

The Fc portion of the molecule was generated in a PCR reaction with pFc-A3 using the primers

1216-52	AAC ATA AGT ACC TGT AGG ATC G
1798-17	AGA GTA AGT ACC TCC ACC ACC ACC TCC ACC TTT ACC CGG AGA CAG GGA GAG GCT CTT CTG C

which are SEQ ID NOS: 398 and 399, respectively. The oligonucleotides 1798-17 and 1798-18 contain an overlap of 61 nucleotides, allowing the two genes to be fused together in the correct reading frame by combining the above PCR products in a third reaction using the outside primers, 1216-52 and 1798-19.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 (described below), also digested with XbaI and BamHI. Ligated DNA was transformed into competent host cells of E. coli strain 2596 (GM221, described herein). Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3718.

The nucleotide and amino acid sequence of the resulting fusion protein (SEQ ID NOS: 15 and 16) are shown in Figure 13.

EMP-Fc. A DNA sequence coding for a monomer of the EPO-mimetic peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. Templates for PCR reactions were the pFC-A3a vector and a synthetic gene encoding EPO monomer. The synthetic gene for the monomer was constructed from the 4 overlapping oligonucleotides 1798-4 and 1798-5 (above) and 1798-6 and 1798-7 (SEQ ID NOS: 400 and 401, respectively) shown below:

1798-6 GGC CCG CTG ACC TGG GTA TGT AAG CCA CAA GGG GGT GGG
GGA GGC GGG GGG TAA TCT CGA G

1798-7 GAT CCT CGA GAT TAC CCC CCG CCT CCC CCA CCC CCT TGT
GGC TTA CAT AC

The 4 oligonucleotides were annealed to form the duplex encoding an amino acid sequence (SEQ ID NOS: 402 and 403, respectively) shown below:

```
          GTTTGCAAACCGCAGGGTGGCGGCGGCGGCGGCGGTGGTACCTATTCCTGTCATTTTGGC
        1 ---------+---------+---------+---------+---------+---------+ 60
                   GTCCCACCGCCGCCGCCGCCGCCACCATGGATAAGGACAGTAAAACCG
A         V  C  K  P  Q  G  G  G  G  G  G  G  G  T  Y  S  C  H  F  G   -

          CCGCTGACCTGGGTATGTAAGCCACAAGGGGGTGGGGGAGGCGGGGGGTAATCTCGAG
       61 ---------+---------+---------+---------+---------+---------+- 122
          GGCGACTGGACCCATACATTCGGTGTTCCCCCACCCCCTCCGCCCCCCATTAGAGCTCCTAG
A         P  L  T  W  V  C  K  P  Q  G  G  G  G  G  G  G  *
```

This duplex was amplified in a PCR reaction using

1798-21 TTA TTT CAT ATG AAA GGT GGT AAC TAT TCC TGT CAT TTT

and

1798-22 TGG ACA TGT GTG AGT TTT GTC CCC CCC GCC TCC CCC ACC
CCC T

as the sense and antisense primers (SEQ ID NOS: 404 and 405, respectively).

The Fc portion of the molecule was generated in a PCR reaction with pFc-A3 using the primers

1798-23 AGG GGG TGG GGG AGG CGG GGG GGA CAA AAC TCA CAC ATG
TCC A

and

1200-54 GTT ATT GCT CAG CGG TGG CA

which are SEQ ID NOS: 406 and 407, respectively. The oligonucleotides 1798-22 and 1798-23 contain an overlap of 43 nucleotides, allowing the two genes to be fused together in the correct reading frame by combining the above PCR products in a third reaction using the outside primers, 1787-21 and 1200-54.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated

into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described above. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3688.

The nucleotide and amino acid sequences (SEQ ID NOS: 17 and 18) of the resulting fusion protein are shown in Figure 14.

EMP-EMP-Fc. A DNA sequence coding for a dimer of the EPO-mimetic peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. Templates for PCR reactions were the EMP-Fc plasmid from strain #3688 above and a synthetic gene encoding the EPO dimer. The synthetic gene for the dimer was constructed from the 8 overlapping oligonucleotides (SEQ ID NOS:408 to 415, respectively) shown below:

1869-23	TTT TTT ATC GAT TTG ATT CTA GAT TTG AGT TTT AAC TTT TAG AAG GAG GAA TAA AAT ATG
1869-48	TAA AAG TTA AAA CTC AAA TCT AGA ATC AAA TCG ATA AAA AA
1871-72	GGA GGT ACT TAC TCT TGC CAC TTC GGC CCG CTG ACT TGG GTT TGC AAA CCG
1871-73	AGT CAG CGG GCC GAA GTG GCA AGA GTA AGT ACC TCC CAT ATT TTA TTC CTC CTT C
1871-74	CAG GGT GGC GGC GGC GGC GGC GGT GGT ACC TAT TCC TGT CAT TTT GGC CCG CTG ACC TGG
1871-75	AAA ATG ACA GGA ATA GGT ACC ACC GCC GCC GCC GCC GCC ACC CTG CGG TTT GCA AAC CCA
1871-78	GTA TGT AAG CCA CAA GGG GGT GGG GGA GGC GGG GGG GAC AAA ACT CAC ACA TGT CCA
1871-79	AGT TTT GTC CCC CCC GCC TCC CCC ACC CCC TTG TGG CTT ACA TAC CCA GGT CAG CGG GCC

The 8 oligonucleotides were annealed to form the duplex encoding an amino acid sequence (SEQ ID NOS: 416 and 417, respectively) shown below:

```
                TTTTTTATCGATTTGATTCTAGATTTGAGTTTTAACTTTTAGAAGGAGGAATAAAATATG
          1     ---------+---------+---------+---------+---------+---------+ 60
                AAAAAATAGCTAAACTAAGATCTAAACTCAAAATTGAAAATCTTCCTCCTTATTTTATAC
a                                                                       M
```

```
                GGAGGTACTTACTCTTGCCACTTCGGCCCGCTGACTTGGGTTTGCAAACCGCAGGGTGGC
           61   ---------+---------+---------+---------+---------+---------+  120
                CCTCCATGAATGAGAACGGTGAAGCCGGGCGACTGAACCCAAACGTTTGGCGTCCCACCG
 5  a             G  G  T  Y  S  C  H  F  G  P  L  T  W  V  C  K  P  Q  G  G   -

                GGCGGCGGCGGCGGTGGTACCTATTCCTGTCATTTTGGCCCGCTGACCTGGGTATGTAAG
          121   ---------+---------+---------+---------+---------+---------+  180
                CCGCCGCCGCCGCCACCATGGATAAGGACAGTAAAACCGGGCGACTGGACCCATACATTC
10  a             G  G  G  G  G  G  T  Y  S  C  H  F  G  P  L  T  W  V  C  K   -

                CCACAAGGGGGTGGGGGAGGCGGGGGGGACAAAACTCACACATGTCCA
          181   ---------+---------+---------+---------+--------  228
                GGTGTTCCCCCACCCCCTCCGCCCCCCCTGTTTTGA
15  a             P  Q  G  G  G  G  G  G  G  D  K  T  H  T  C  P   -
```

This duplex was amplified in a PCR reaction using 1869-23 and 1871-79 (shown above) as the sense and antisense primers.

The Fc portion of the molecule was generated in a PCR reaction with strain 3688 DNA using the primers 1798-23 and 1200-54 (shown above).

The oligonucleotides 1871-79 and 1798-23 contain an overlap of 31 nucleotides, allowing the two genes to be fused together in the correct reading frame by combining the above PCR products in a third reaction using the outside primers, 1869-23 and 1200-54.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for Fc-EMP. Clones were screened for ability to produce the recombinant protein product and possession of the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3813.

The nucleotide and amino acid sequences (SEQ ID NOS: 19 and 20, respectively) of the resulting fusion protein are shown in Figure 15. There is a silent mutation at position 145 (A to G, shown in boldface) such that the final construct has a different nucleotide sequence than the oligonucleotide 1871-72 from which it was derived.

Fc-EMP-EMP. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a dimer of the EPO-mimetic peptide was

constructed using standard PCR technology. Templates for PCR reactions were the plasmids from strains 3688 and 3813 above.

The Fc portion of the molecule was generated in a PCR reaction with strain 3688 DNA using the primers 1216-52 and 1798-17 (shown above). The EMP dimer portion of the molecule was the product of a second PCR reaction with strain 3813 DNA using the primers 1798-18 (also shown above) and SEQ ID NO: 418, shown below:

1798-20 CTA ATT GGA TCC TCG AGA TTA ACC CCC TTG TGG CTT ACAT

The oligonucleotides 1798-17 and 1798-18 contain an overlap of 61 nucleotides, allowing the two genes to be fused together in the correct reading frame by combining the above PCR products in a third reaction using the outside primers, 1216-52 and 1798-20.

The final PCR gene product (the full length fusion gene) was digested with restriction endonucleases XbaI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for Fc-EMP. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #3822.

The nucleotide and amino acid sequences (SEQ ID NOS: __ and __, respectively) of the fusion protein are shown in Figure 16.

Characterization of Fc-EMP activity. Characterization was carried out in vivo as follows.

Mice: Normal female BDF1 approximately 10-12 weeks of age.

Bleed schedule: Ten mice per group treated on day 0, two groups started 4 days apart for a total of 20 mice per group. Five mice bled at each time point, mice were bled a maximum of three times a week. Mice were anesthetized with isoflurane and a total volume of 140-160 ml of blood was obtained by puncture of the orbital sinus. Blood was counted

on a Technicon H1E blood analyzer running software for murine blood. Parameters measured were WBC, RBC, HCT, HGB, PLT, NEUT, LYMPH.

Treatments: Mice were either injected subcutaneously for a bolus treatment or implanted with 7 day micro-osmotic pumps for continuous delivery. Subcutaneous injections were delivered in a volume of 0.2 ml. Osmotic pumps were inserted into a subcutaneous incision made in the skin between the scapulae of anesthetized mice. Compounds were diluted in PBS with 0.1% BSA. All experiments included one control group, labeled "carrier" that were treated with this diluent only. The concentration of the test articles in the pumps was adjusted so that the calibrated flow rate from the pumps gave the treatment levels indicated in the graphs.

Experiments: Various Fc-conjugated EPO mimetic peptides (EMPs) were delivered to mice as a single bolus injection at a dose of 100 μg/kg. Fc-EMPs were delivered to mice in 7-day micro-osmotic pumps. The pumps were not replaced at the end of 7 days. Mice were bled until day 51 when HGB and HCT returned to baseline levels.

Example 4

TNF-α inhibitors

Fc-TNF-α inhibitors. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of the TNF-α inhibitory peptide was constructed using standard PCR technology. The Fc and 5 glycine linker portion of the molecule was generated in a PCR reaction with DNA from the Fc-EMP fusion strain #3718 (see Example 3) using the sense primer 1216-52 and the antisense primer 2295-89 (SEQ ID NOS: 1112 and 1113 , respectively). The nucleotides encoding the TNF-α inhibitory peptide were provided by the PCR primer 2295-89 shown below:

1216-52 AAC ATA AGT ACC TGT AGG ATC G

2295-89 CCG CGG ATC CAT TAC GGA CGG TGA CCC AGA GAG GTG TTT TTG TAG

TGC GGC AGG AAG TCA CCA CCA CCT CCA CCT TTA CCC

The oligonucleotide 2295-89 overlaps the glycine linker and Fc portion of the template by 22 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4544.

The nucleotide and amino acid sequences (SEQ ID NOS: 1055 and 1056) of the fusion protein are shown in Figures 19A and 19B.

TNF-α inhibitor-Fc. A DNA sequence coding for a TNF-α inhibitory peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. The template for the PCR reaction was a plasmid containing an unrelated peptide fused via a five glycine linker to Fc. The nucleotides encoding the TNF-α inhibitory peptide were provided by the sense PCR primer 2295-88, with primer 1200-54 serving as the antisense primer (SEQ ID NOS: 1117 and 407, respectively). The primer sequences are shown below:

2295-88 GAA TAA CAT ATG GAC TTC CTG CCG CAC TAC AAA AAC ACC TCT CTG GGT CAC CGT CCG GGT GGA GGC GGT GGG GAC AAA ACT

1200-54 GTT ATT GCT CAG CGG TGG CA

The oligonucleotide 2295-88 overlaps the glycine linker and Fc portion of the template by 24 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4543.

The nucleotide and amino acid sequences (SEQ ID NOS: 1057 and 1058) of the fusion protein are shown in Figures 20A and 20B.

Expression in E. coli. Cultures of each of the pAMG21-Fc-fusion constructs in E. coli GM221 were grown at 37 °C in Luria Broth medium containing 50 mg/ml kanamycin. Induction of gene product expression from the luxPR promoter was achieved following the addition of the synthetic autoinducer N-(3-oxohexanoyl)-DL-homoserine lactone to the culture media to a final concentration of 20 ng/ml. Cultures were incubated at 37 °C for a further 3 hours. After 3 hours, the bacterial cultures were examined by microscopy for the presence of inclusion bodies and were then collected by centrifugation. Refractile inclusion bodies were observed in induced cultures indicating that the Fc-fusions were most likely produced in the insoluble fraction in E. coli. Cell pellets were lysed directly by resuspension in Laemmli sample buffer containing 10% β-mercaptoethanol and were analyzed by SDS-PAGE. In each case, an intense coomassie-stained band of the appropriate molecular weight was observed on an SDS-PAGE gel.

Purification of Fc-peptide fusion proteins. Cells are broken in water (1/10) by high pressure homogenization (2 passes at 14,000 PSI) and inclusion bodies are harvested by centrifugation (4200 RPM in J-6B for 1 hour). Inclusion bodies are solubilized in 6M guanidine, 50mM Tris, 8mM DTT, pH 8.7 for 1 hour at a 1/10 ratio. The solubilized mixture is diluted

20 times into 2M urea, 50 mM tris, 160mM arginine, 3mM cysteine, pH 8.5. The mixture is stirred overnight in the cold and then concentrated about 10 fold by ultafiltration. It is then diluted 3 fold with 10mM Tris, 1.5M urea, pH 9. The pH of this mixture is then adjusted to pH 5 with acetic acid. The precipitate is removed by centrifugation and the supernatant is loaded onto a SP-Sepharose Fast Flow column equilibrated in 20mM NaAc, 100 mM NaCl, pH 5 (10mg/ml protein load, room temperature). The protein is eluted from the column using a 20 column volume gradient in the same buffer ranging from 100mM NaCl to 500mM NaCl. The pool from the column is diluted 3 fold and loaded onto a SP-Sepharose HP column in 20mM NaAc, 150mM NaCl, pH 5(10mg/ml protein load, room temperature). The protein is eluted using a 20 column volume gradient in the same buffer ranging from 150mM NaCl to 400mM NaCl. The peak is pooled and filtered.

Characterization of activity of Fc-TNF-α inhibitor and TNF-α inhibitor -Fc. Binding of these peptide fusion proteins to TNF- α can be characterized by BIAcore by methods available to one of ordinary skill in the art who is armed with the teachings of the present specification.

Example 5

IL-1 Antagonists

Fc-IL-1 antagonist. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of an IL-1 antagonist peptide was constructed using standard PCR technology. The Fc and 5 glycine linker portion of the molecule was generated in a PCR reaction with DNA from the Fc-EMP fusion strain #3718 (see Example 3) using the sense primer 1216-52 and the antisense primer 2269-70 (SEQ ID NOS: 1112 and 1118, respectively). The nucleotides encoding the IL-1 antagonist peptide were provided by the PCR primer 2269-70 shown below:

1216-52 AAC ATA AGT ACC TGT AGG ATC G

2269-70 CCG CGG ATC CAT TAC AGC GGC AGA GCG TAC GGC TGC CAG TAA CCC
GGG GTC CAT TCG AAA CCA CCA CCT CCA CCT TTA CCC

The oligonucleotide 2269-70 overlaps the glycine linker and Fc portion of the template by 22 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4506.

The nucleotide and amino acid sequences (SEQ ID NOS: 1059 and 1060) of the fusion protein are shown in Figures 21A and 21B.

IL-1 antagonist-Fc. A DNA sequence coding for an IL-1 antagonist peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. The template for the PCR reaction was a plasmid containing an unrelated peptide fused via a five glycine linker to Fc. The nucleotides encoding the IL-1 antagonist peptide were provided by the sense PCR primer 2269-69, with primer 1200-54 serving as the antisense primer (SEQ ID NOS: 1119 and 407, respectively). The primer sequences are shown below:

2269-69 GAA TAA CAT ATG TTC GAA TGG ACC CCG GGT TAC TGG CAG CCG TAC GCT
CTG CCG CTG GGT GGA GGC GGT GGG GAC AAA ACT

1200-54 GTT ATT GCT CAG CGG TGG CA

The oligonucleotide 2269-69 overlaps the glycine linker and Fc portion of the template by 24 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4505.

The nucleotide and amino acid sequences (SEQ ID NOS: 1061 and 1062) of the fusion protein are shown in Figures 22A and 22B. Expression and purification were carried out as in previous examples.

Characterization of Fc-IL-1 antagonist peptide and IL-1 antagonist peptide-Fc activity. IL-1 Receptor Binding competition between IL-1β, IL-1RA and Fc-conjugated IL-1 peptide sequences was carried out using the IGEN system. Reactions contained 0.4 nM biotin-IL-1R + 15 nM IL-1-TAG + 3 uM competitor + 20 ug/ml streptavidin-conjugate beads, where competitors were IL-1RA, Fc-IL-1 antagonist, IL-1 antagonist-Fc). Competition was assayed over a range of competitor concentrations from 3 uM to 1.5 pM. The results are shown in Table C below:

Table C—Results from IL-1 Receptor Binding Competition Assay

	IL-1pep-Fc	*Fc-IL-1pep*	*IL-1ra*
KI	281.5	59.58	1.405
EC50	530.0	112.2	2.645
95% Confidence Intervals			
EC50	280.2 to 1002	54.75 to 229.8	1.149 to 6.086
KI	148.9 to 532.5	29.08 to 122.1	0.6106 to 3.233
Goodness of Fit			
R^2	0.9790	0.9687	0.9602

Example 6

VEGF-Antagonists

Fc-VEGF Antagonist. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of the VEGF mimetic peptide was constructed using standard PCR technology. The templates for the PCR reaction were the pFc-A3 plasmid and a synthetic VEGF mimetic peptide gene. The synthetic gene was assembled by annealing the following two oligonucleotides primer (SEQ ID NOS: 1120 and 1121, respectively):

2293-11 GTT GAA CCG AAC TGT GAC ATC CAT GTT ATG TGG GAA TGG GAA TGT TTT GAA CGT CTG

2293-12 CAG ACG TTC AAA ACA TTC CCA TTC CCA CAT AAC ATG GAT GTC ACA GTT CGG TTC AAC

The two oligonucleotides anneal to form the following duplex encoding an amino acid sequence shown below (SEQ ID NOS 1122):

```
                GTTGAACCGAACTGTGACATCCATGTTATGTGGGAATGGGAATGTTTTGAACGTCTG
             1 ---------+---------+---------+---------+---------+------- 57
                CAACTTGGCTTGACACTGTAGGTACAATACACCCTTACCCTTACAAAACTTGCAGAC

a               V  E  P  N  C  D  I  H  V  M  W  E  W  E  C  F  E  R  L
```

This duplex was amplified in a PCR reaction using 2293-05 and 2293-06 as the sense and antisense primers (SEQ ID NOS. 1125 and 1126).

The Fc portion of the molecule was generated in a PCR reaction with the pFc-A3 plasmid using the primers 2293-03 and 2293-04 as the sense and antisense primers (SEQ ID NOS. 1123 and 1124, respectively). The full length fusion gene was obtained from a third PCR reaction using the outside primers 2293-03 and 2293-06. These primers are shown below:

2293-03 ATT TGA TTC TAG AAG GAG GAA TAA CAT ATG GAC AAA ACT CAC ACA TGT

2293-04 GTC ACA GTT CGG TTC AAC ACC ACC ACC ACC ACC TTT ACC CGG AGA CAG GGA

2293-05 TCC CTG TCT CCG GGT AAA GGT GGT GGT GGT GGT GTT GAA CCG AAC TGT GAC ATC

2293-06 CCG CGG ATC CTC GAG TTA CAG ACG TTC AAA ACA TTC CCA

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4523.

The nucleotide and amino acid sequences (SEQ ID NOS: 1063 and 1064) of the fusion protein are shown in Figures 23A and 23B.

VEGF antagonist -Fc. A DNA sequence coding for a VEGF mimetic peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. The templates for the PCR reaction were the pFc-A3 plasmid and the synthetic VEGF mimetic peptide gene described above. The synthetic duplex was amplified in a PCR reaction using 2293-07 and 2293-08 as the sense and antisense primers (SEQ ID NOS. 1127 and 1128, respectively).

The Fc portion of the molecule was generated in a PCR reaction with the pFc-A3 plasmid using the primers 2293-09 and 2293-10 as the sense and antisense primers (SEQ ID NOS. 1129 and 1130, respectively).

The full length fusion gene was obtained from a third PCR reaction using the outside primers 2293-07 and 2293-10. These primers are shown below:

2293-07 ATT TGA TTC TAG AAG GAG GAA TAA CAT ATG GTT GAA CCG AAC TGT GAC

2293-08 ACA TGT GTG AGT TTT GTC ACC ACC ACC ACC ACC CAG ACG TTC AAA ACA TTC

2293-09 GAA TGT TTT GAA CGT CTG GGT GGT GGT GGT GGT GAC AAA ACT CAC ACA TGT

2293-10 CCG CGG ATC CTC GAG TTA TTT ACC CGG AGA CAG GGA GAG

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4524.

The nucleotide and amino acid sequences (SEQ ID NOS: 1065 and 1066) of the fusion protein are shown in Figures 24A and 24B. Expression and purification were carried out as in previous examples.

Example 7

MMP Inhibitors

Fc-MMP inhibitor. A DNA sequence coding for the Fc region of human IgG1 fused in-frame to a monomer of an MMP inhibitory peptide was constructed using standard PCR technology. The Fc and 5 glycine linker portion of the molecule was generated in a PCR reaction with DNA from the Fc-TNF-α inhibitor fusion strain #4544 (see Example 4) using the sense primer 1216-52 and the antisense primer 2308-67 (SEQ ID NOS: 1112

and 1131, respectively). The nucleotides encoding the MMP inhibitor peptide were provided by the PCR primer 2308-67 shown below:

1216-52 AAC ATA AGT ACC TGT AGG ATC G

2308-67 CCG CGG ATC CAT TAG CAC AGG GTG AAA CCC CAG TGG GTG GTG CAA CCA CCA CCT CCA CCT TTA CCC

The oligonucleotide 2308-67 overlaps the glycine linker and Fc portion of the template by 22 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4597.

The nucleotide and amino acid sequences (SEQ ID NOS: 1067 and 1068) of the fusion protein are shown in Figures 25A and 25B. Expression and purification were carried out as in previous examples.

MMP Inhibitor-Fc. A DNA sequence coding for an MMP inhibitory peptide fused in-frame to the Fc region of human IgG1 was constructed using standard PCR technology. The Fc and 5 glycine linker portion of the molecule was generated in a PCR reaction with DNA from the Fc-TNF-α inhibitor fusion strain #4543 (see Example 4). The nucleotides encoding the MMP inhibitory peptide were provided by the sense PCR primer 2308-66, with primer 1200-54 serving as the antisense primer (SEQ ID NOS: 1132 and 407, respectively). The primer sequences are shown below:

2308-66 GAA TAA CAT ATG TGC ACC ACC CAC TGG GGT TTC ACC CTG TGC GGT GGA GGC GGT GGG GAC AAA

1200-54 GTT ATT GCT CAG CGG TGG CA

The oligonucleotide 2269-69 overlaps the glycine linker and Fc portion of the template by 24 nucleotides, with the PCR resulting in the two genes being fused together in the correct reading frame.

The PCR gene product (the full length fusion gene) was digested with restriction endonucleases NdeI and BamHI, and then ligated into the vector pAMG21 and transformed into competent E. coli strain 2596 cells as described for EMP-Fc herein. Clones were screened for the ability to produce the recombinant protein product and to possess the gene fusion having the correct nucleotide sequence. A single such clone was selected and designated Amgen strain #4598.

The nucleotide and amino acid sequences (SEQ ID NOS: 1069 and 1070) of the fusion protein are shown in Figures 26A and 26B.

* * *

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto, without departing from the spirit and scope of the invention as set forth herein.

Abbreviations

Abbreviations used throughout this specification are as defined below, unless otherwise defined in specific circumstances.

Ac	acetyl (used to refer to acetylated residues)
AcBpa	acetylated p-benzoyl-L-phenylalanine
ADCC	antibody-dependent cellular cytotoxicity
Aib	aminoisobutyric acid
bA	beta-alanine
Bpa	p-benzoyl-L-phenylalanine
BrAc	bromoacetyl ($BrCH_2C(O)$

BSA	Bovine serum albumin
Bzl	Benzyl
Cap	Caproic acid
CTL	Cytotoxic T lymphocytes
CTLA4	Cytotoxic T lymphocyte antigen 4
DARC	Duffy blood group antigen receptor
DCC	Dicylcohexylcarbodiimide
Dde	1-(4,4-dimethyl-2,6-dioxo-cyclohexylidene)ethyl
EMP	Erythropoietin-mimetic peptide
ESI-MS	Electron spray ionization mass spectrometry
EPO	Erythropoietin
Fmoc	fluorenylmethoxycarbonyl
G-CSF	Granulocyte colony stimulating factor
GH	Growth hormone
HCT	hematocrit
HGB	hemoglobin
hGH	Human growth hormone
HOBt	1-Hydroxybenzotriazole
HPLC	high performance liquid chromatography
IL	interleukin
IL-R	interleukin receptor
IL-1R	interleukin-1 receptor
IL-1ra	interleukin-1 receptor antagonist
Lau	Lauric acid
LPS	lipopolysaccharide
LYMPH	lymphocytes
MALDI-MS	Matrix-assisted laser desorption ionization mass spectrometry
Me	methyl

MeO	methoxy
MHC	major histocompatibility complex
MMP	matrix metalloproteinase
MMPI	matrix metalloproteinase inhibitor
1-Nap	1-napthylalanine
NEUT	neutrophils
NGF	nerve growth factor
Nle	norleucine
NMP	N-methyl-2-pyrrolidinone
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
Pbf	2,2,4,6,7-pendamethyldihydrobenzofuran-5-sulfonyl
PCR	polymerase chain reaction
Pec	pipecolic acid
PEG	Poly(ethylene glycol)
pGlu	pyroglutamic acid
Pic	picolinic acid
PLT	platelets
pY	phosphotyrosine
RBC	red blood cells
RBS	ribosome binding site
RT	room temperature (25 °C)
Sar	sarcosine
SDS	sodium dodecyl sulfate
STK	serine-threonine kinases
t-Boc	tert-Butoxycarbonyl
tBu	tert-Butyl
TGF	tissue growth factor
THF	thymic humoral factor

TK	tyrosine kinase
TMP	Thrombopoietin-mimetic peptide
TNF	Tissue necrosis factor
TPO	Thrombopoietin
TRAIL	TNF-related apoptosis-inducing ligand
Trt	trityl
UK	urokinase
UKR	urokinase receptor
VEGF	vascular endothelial cell growth factor
VIP	vasoactive intestinal peptide
WBC	white blood cells

What is claimed is:

1. A composition of matter of the formula

$$(X^1)_a\text{-}F^1\text{-}(X^2)_b$$

and multimers thereof, wherein:

F^1 is an Fc domain;

X^1 and X^2 are each independently selected from $-(L^1)_c\text{-}P^1$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3$, and $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3\text{-}(L^4)_f\text{-}P^4$

P^1, P^2, P^3, and P^4 are each independently sequences of pharmacologically active peptides;

L^1, L^2, L^3, and L^4 are each independently linkers; and

a, b, c, d, e, and f are each independently 0 or 1, provided that at least one of a and b is 1.

2. The composition of matter of Claim 1 of the formulae

$$X^1\text{-}F^1$$

or

$$F^1\text{-}X^2.$$

3. The composition of matter of Claim 1 of the formula

$$F^1\text{-}(L^1)_c\text{-}P^1.$$

4. The composition of matter of Claim 1 of the formula

$$F^1\text{-}(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2.$$

5. The composition of matter of Claim 1 wherein F^1 is an IgG Fc domain.

6. The composition of matter of Claim 1 wherein F^1 is an IgG1 Fc domain.

7. The composition of matter of Claim 1 wherein F^1 comprises the sequence of SEQ ID NO: 2.

8. The composition of matter of Claim 1 wherein X^1 and X^2 comprise an IL-1 antagonist peptide sequence.

9. The composition of matter of Claim 8 wherein the IL-1 antagonist peptide sequence is selected from SEQ ID NOS: 212, 907, 908, 909, 910, 917, and 979.

10. The composition of matter of Claim 8 wherein the IL-1 antagonist peptide sequence is selected from SEQ ID NOS: 213 to 271, 671 to 906, 911 to 916, and 918 to 1023.

11. The composition of matter of Claim 8 wherein F^1 comprises the sequence of SEQ ID NO: 2.

12. The composition of matter of Claim 1 wherein X^1 and X^2 comprise an EPO-mimetic peptide sequence.

13. The composition of matter of Claim 12 wherein the EPO-mimetic peptide sequence is selected from Table 5.

14. The composition of matter of Claim 12 wherein F^1 comprises the sequence of SEQ ID NO: 2.

15. The composition of matter of Claim 12 comprising a sequence selected from SEQ ID NOS: 83, 84, 85, 124, 419, 420, 421, and 461. .

16. The composition of matter of claim 12 comprising a sequence selected from SEQ ID NOS: 339 and 340.

17. The composition of matter of Claim 12 comprising a sequence selected from SEQ ID NOS: 20 and 22.

18. The composition of matter of Claim 3 wherein P^1 is a TPO-mimetic peptide sequence.

19. The composition of matter of Claim 18 wherein P^1 is a TPO-mimetic peptide sequence selected from Table 6.

20. The composition of matter of Claim 18 wherein F^1 comprises the sequence of SEQ ID NO: 2.

21. The composition of matter of Claim 18 having a sequence selected from SEQ ID NOS: 6 and 12.

22. A DNA encoding a composition of matter of any of Claims 1 to 21.

23. An expression vector comprising the DNA of Claim 22.

24. A host cell comprising the expression vector of Claim 23.

25. The cell of Claim 24, wherein the cell is an E. coli cell.

26. A process for preparing a pharmacologically active compound, which comprises

a) selecting at least one randomized peptide that modulates the activity of a protein of interest; and

b) preparing a pharmacologic agent comprising at least one Fc domain covalently linked to at least one amino acid sequence of the selected peptide or peptides.

27. The process of Claim 26, wherein the peptide is selected in a process comprising screening of a phage display library, an E. coli display library, a ribosomal library, or a chemical peptide library.

28. The process of Claim 26, wherein the preparation of the pharmacologic agent is carried out by:

a) preparing a gene construct comprising a nucleic acid sequence encoding the selected peptide and a nucleic acid sequence encoding an Fc domain; and

b) expressing the gene construct.

29. The process of Claim 26, wherein the gene construct is expressed in an E. coli cell.

30. The process of Claim 26, wherein the protein of interest is a cell surface receptor.

31. The process of Claim 26, wherein the protein of interest has a linear epitope.

32. The process of Claim 26, wherein the protein of interest is a cytokine receptor.

33. The process of Claim 26, wherein the peptide is an EPO-mimetic peptide.

34. The process of Claim 26, wherein the peptide is a TPO-mimetic peptide.

35. The process of Claim 26, wherein the peptide is an IL-1 antagonist peptide.

36. The process of Claim 26, wherein the peptide is an MMP inhibitor peptide or a VEGF antagonist peptide.

37. The process of Claim 26, wherein the peptide is a TNF-antagonist peptide.

38. The process of Claim 26, wherein the peptide is a CTLA4-mimetic peptide.

39. The process of Claim 26, wherein the peptide is selected from Tables 4 to 20.

40. The process of Claim 26, wherein the selection of the peptide is carried out by a process comprising:

a) preparing a gene construct comprising a nucleic acid sequence encoding a first selected peptide and a nucleic acid sequence encoding an Fc domain;

b) conducting a polymerase chain reaction using the gene construct and mutagenic primers, wherein

i) a first mutagenic primer comprises a nucleic acid sequence complementary to a sequence at or near the 5′ end of a coding strand of the gene construct, and

ii) a second mutagenic primer comprises a nucleic acid sequence complementary to the 3′ end of the noncoding strand of the gene construct.

41. The process of Claim 26, wherein the compound is derivatized.

42. The process of Claim 26, wherein the derivatized compound comprises a cyclic portion, a cross-linking site, a non-peptidyl

linkage, an N-terminal replacement, a C-terminal replacement, or a modified amino acid moiety.

43. The process of Claim 26 wherein the Fc domain is an IgG Fc domain.

44. The process of Claim 26, wherein the vehicle is an IgG1 Fc domain.

45. The process of Claim 26, wherein the vehicle comprises the sequence of SEQ ID NO: 2.

46. The process of Claim 26, wherein the compound prepared is of the formula

$$(X^1)_a\text{-}F^1\text{-}(X^2)_b$$

and multimers thereof,wherein:

F^1 is an Fc domain;

X^1 and X^2 are each independently selected from $-(L^1)_c\text{-}P^1$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2$, $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3$, and $-(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2\text{-}(L^3)_e\text{-}P^3\text{-}(L^4)_f\text{-}P^4$

P^1, P^2, P^3, and P^4 are each independently sequences of pharmacologically active peptides;

L^1, L^2, L^3, and L^4 are each independently linkers; and

a, b, c, d, e, and f are each independently 0 or 1, provided that at least one of a and b is 1.

47. The process of Claim 46, wherein the compound prepared is of the formulae

$$X^1\text{-}F^1$$

or

$$F^1\text{-}X^2.$$

48. The process of Claim 46, wherein the compound prepared is of the formulae

$$F^1\text{-}(L^1)_c\text{-}P^1$$

or

$$F^1\text{-}(L^1)_c\text{-}P^1\text{-}(L^2)_d\text{-}P^2.$$

49. The process of Claim 46, wherein F^1 is an IgG Fc domain.

50. The process of Claim 46, wherein F^1 is an IgG1 Fc domain.

51. The process of Claim 46, wherein F^1 comprises the sequence of SEQ ID NO: 2.

FIG. 1

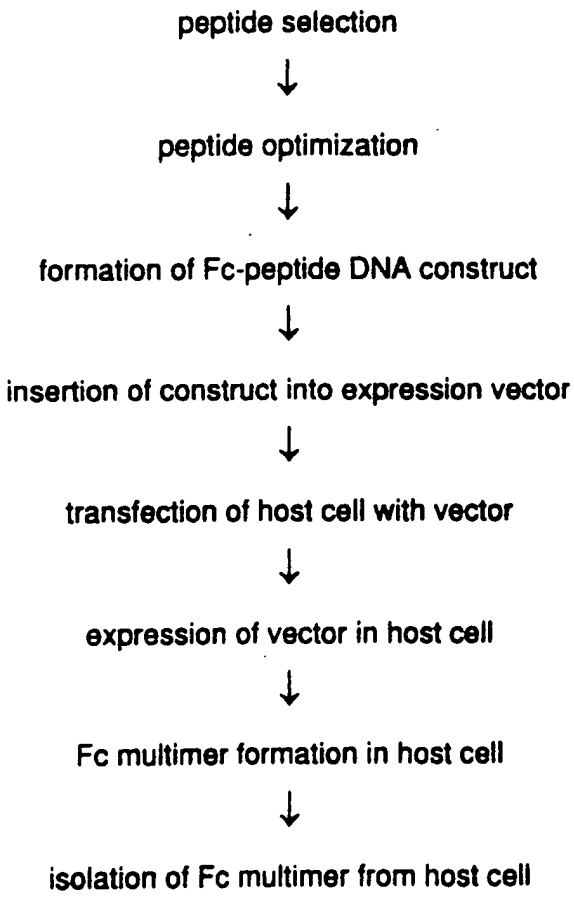

peptide selection
peptide optimization
formation of Fc-peptide DNA construct
insertion of construct into expression vector
transfection of host cell with vector
expression of vector in host cell
Fc multimer formation in host cell
isolation of Fc multimer from host cell

FIG. 2A

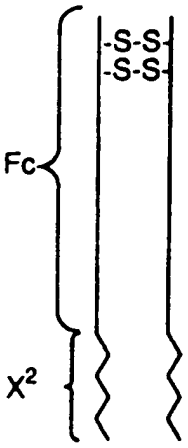


FIG. 2B

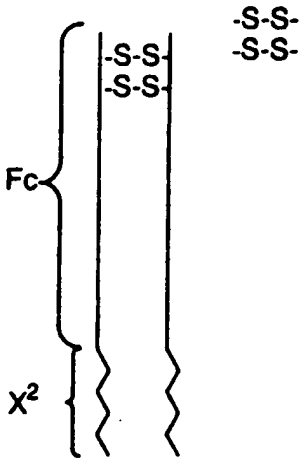


FIG. 2C

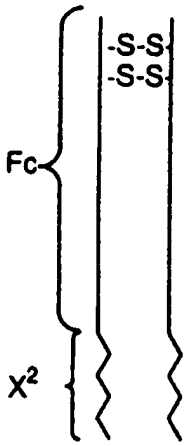


FIG. 2D

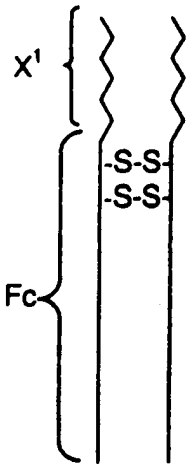


FIG. 2E

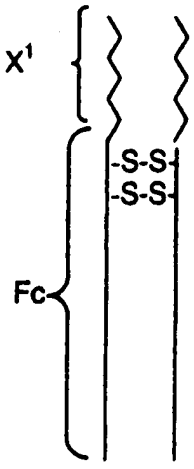


FIG. 2F

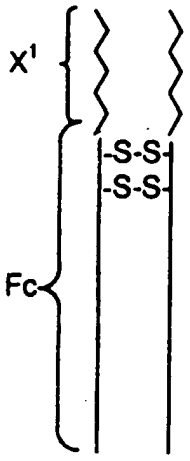


FIG. 3A

FIG. 3B

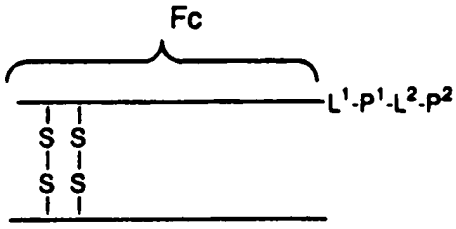


FIG. 3C

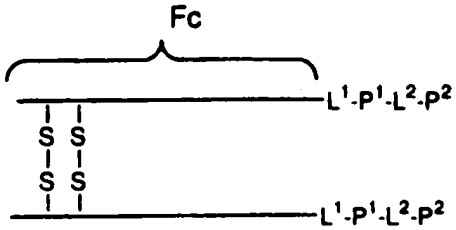


FIG. 4

```
                ATGGACAAAACTCACACATGTCCACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCGTCA
              1 ---------+---------+---------+---------+---------+---------+ 60
                TACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGCAGT

a               M  D  K  T  H  T  C  P  P  C  P  A  P  E  L  L  G  G  P  S  -

                GTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTC
             61 ---------+---------+---------+---------+---------+---------+ 120
                CAGAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAG

a               V  F  L  F  P  P  K  P  K  D  T  L  M  I  S  R  T  P  E  V  -

                ACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTG
            121 ---------+---------+---------+---------+---------+---------+ 180
                TGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCAC

a               T  C  V  V  V  D  V  S  H  E  D  P  E  V  K  F  N  W  Y  V  -

                GACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACG
            181 ---------+---------+---------+---------+---------+---------+ 240
                CTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGC

a               D  G  V  E  V  H  N  A  K  T  K  P  R  E  E  Q  Y  N  S  T  -

                TACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTAC
            241 ---------+---------+---------+---------+---------+---------+ 300
                ATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTCATG

a               Y  R  V  V  S  V  L  T  V  L  H  Q  D  W  L  N  G  K  E  Y  -

                AAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
            301 ---------+---------+---------+---------+---------+---------+ 360
                TTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGG

a               K  C  K  V  S  N  K  A  L  P  A  P  I  E  K  T  I  S  K  A  -

                AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACC
            361 ---------+---------+---------+---------+---------+---------+ 420
                TTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGG

a               K  G  Q  P  R  E  P  Q  V  Y  T  L  P  P  S  R  D  E  L  T  -

                AAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTG
            421 ---------+---------+---------+---------+---------+---------+ 480
                TTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCAC

a               K  N  Q  V  S  L  T  C  L  V  K  G  F  Y  P  S  D  I  A  V  -

                GAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGAC
            481 ---------+---------+---------+---------+---------+---------+ 540
                CTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTG

a               E  W  E  S  N  G  Q  P  E  N  N  Y  K  T  T  P  P  V  L  D  -

                TCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAG
            541 ---------+---------+---------+---------+---------+---------+ 600
                AGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTC

a               S  D  G  S  F  F  L  Y  S  K  L  T  V  D  K  S  R  W  Q  Q  -

                GGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAG
            601 ---------+---------+---------+---------+---------+---------+ 660
                CCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTC

a               G  N  V  F  S  C  S  V  M  H  E  A  L  H  N  H  Y  T  Q  K  -

                AGCCTCTCCCTGTCTCCGGGTAAA
            661 ---------+---------+---- 684
                TCGGAGAGGGACAGAGGCCCATTT
```

FIG. 5

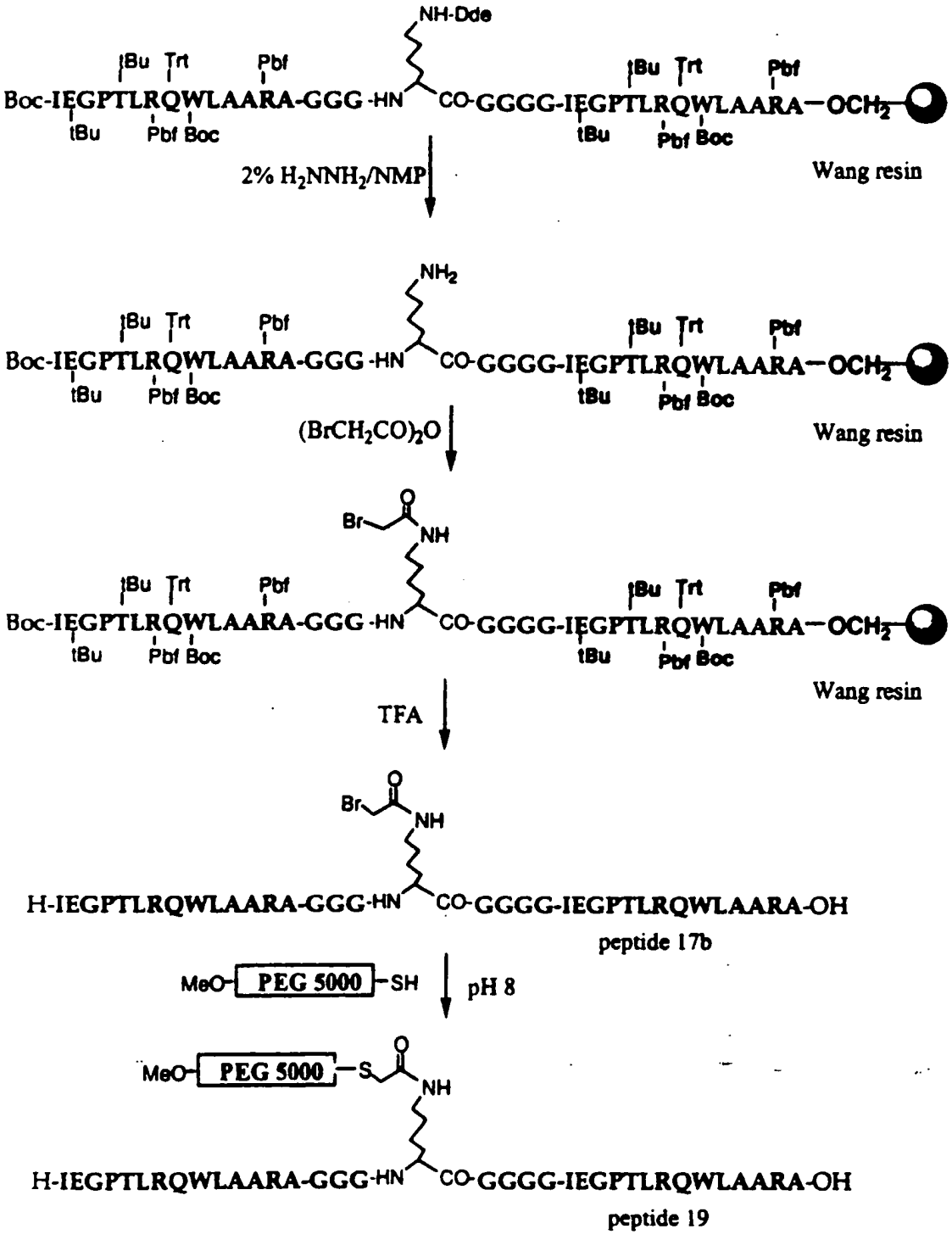

NH-Dde
tBu Trt Pbf
Boc-IEGPTLRQWLAARA-GGG-HN
CO-GGGG-IEGPTLRQWLAARA—OCH2
tBu Pbf Boc
tBu Trt Pbf
tBu Pbf Boc
Wang resin
2% H2NNH2/NMP
NH2
(BrCH2CO)2O
Br
NH
TFA
H-IEGPTLRQWLAARA-GGG-HN
CO-GGGG-IEGPTLRQWLAARA-OH
peptide 17b
MeO-PEG 5000-SH
pH 8
MeO-PEG 5000-S
peptide 19

FIG. 6

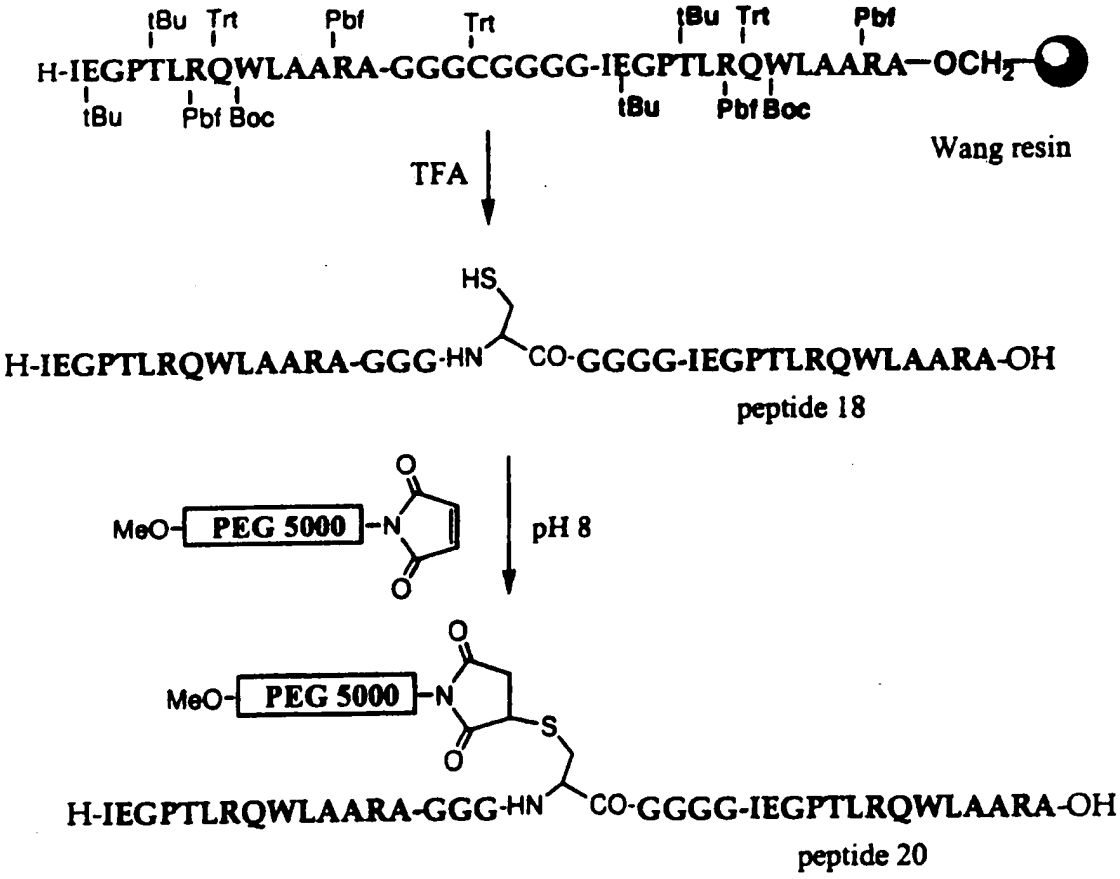

tBu Trt Pbf Trt tBu Trt Pbf
H-IEGPTLRQWLAARA-GGGCGGGG-IEGPTLRQWLAARA—OCH2
tBu Pbf Boc tBu Pbf Boc
Wang resin
TFA
HS
H-IEGPTLRQWLAARA-GGG-HN CO-GGGG-IEGPTLRQWLAARA-OH
peptide 18
MeO- PEG 5000 -N
pH 8
MeO- PEG 5000 -N
S
H-IEGPTLRQWLAARA-GGG-HN CO-GGGG-IEGPTLRQWLAARA-OH
peptide 20

FIG. 7

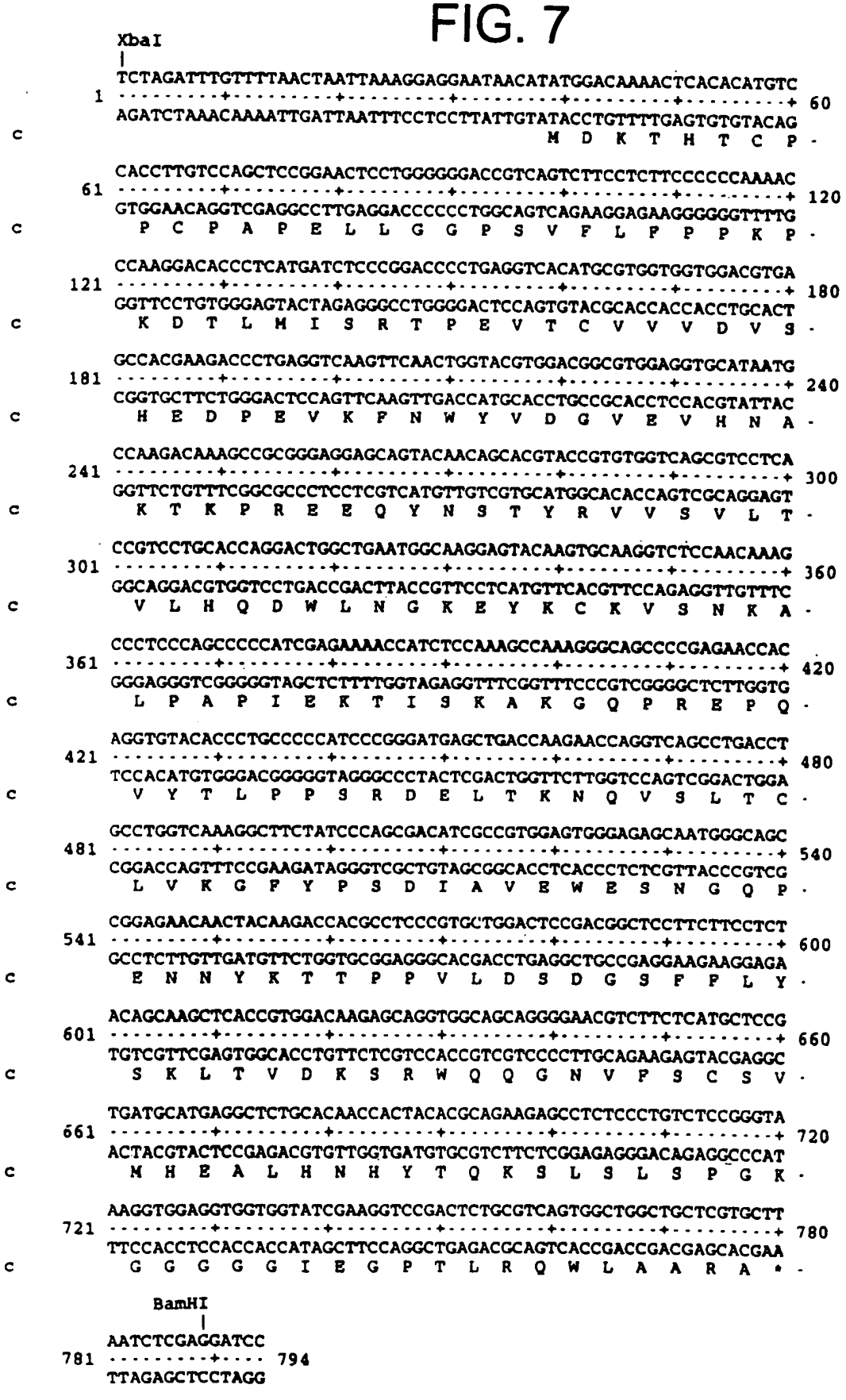

XbaI
TCTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGGACAAAACTCACACATGTC
AGATCTAAACAAAATTGATTAATTTCCTCCTTATTGTATACCTGTTTTGAGTGTGTACAG
c M D K T H T C P -
CACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAAC
GTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGCAGTCAGAAGGAGAAGGGGGGTTTTG
c P C P A P E L L G G P S V F L F P P P K P -
CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGA
GGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACT
c K D T L M I S R T P E V T C V V V D V S -
GCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
CGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTAC
c H E D P E V K F N W Y V D G V E V H N A -
CCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA
GGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCACACCAGTCGCAGGAGT
c K T K P R E E Q Y N S T Y R V V S V L T -
CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAG
GGCAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTC
c V L H Q D W L N G K E Y K C K V S N K A -
CCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCAC
GGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTG
c L P A P I E K T I S K A K G Q P R E P Q -
AGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCT
TCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGA
c V Y T L P P S R D E L T K N Q V S L T C -
GCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGC
CGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCG
c L V K G F Y P S D I A V E W E S N G Q P -
CGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCT
GCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGA
c E N N Y K T T P P V L D S D G S F F L Y -
ACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG
TGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGC
c S K L T V D K S R W Q Q G N V F S C S V -
TGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA
ACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCAT
c M H E A L H N H Y T Q K S L S L S P G K -
AAGGTGGAGGTGGTGGTATCGAAGGTCCGACTCTGCGTCAGTGGCTGGCTGCTCGTGCTT
TTCCACCTCCACCACCATAGCTTCCAGGCTGAGACGCAGTCACCGACCGACGAGCACGAA
c G G G G G I E G P T L R Q W L A A R A * -
BamHI
AATCTCGAGGATCC
TTAGAGCTCCTAGG

FIG. 8

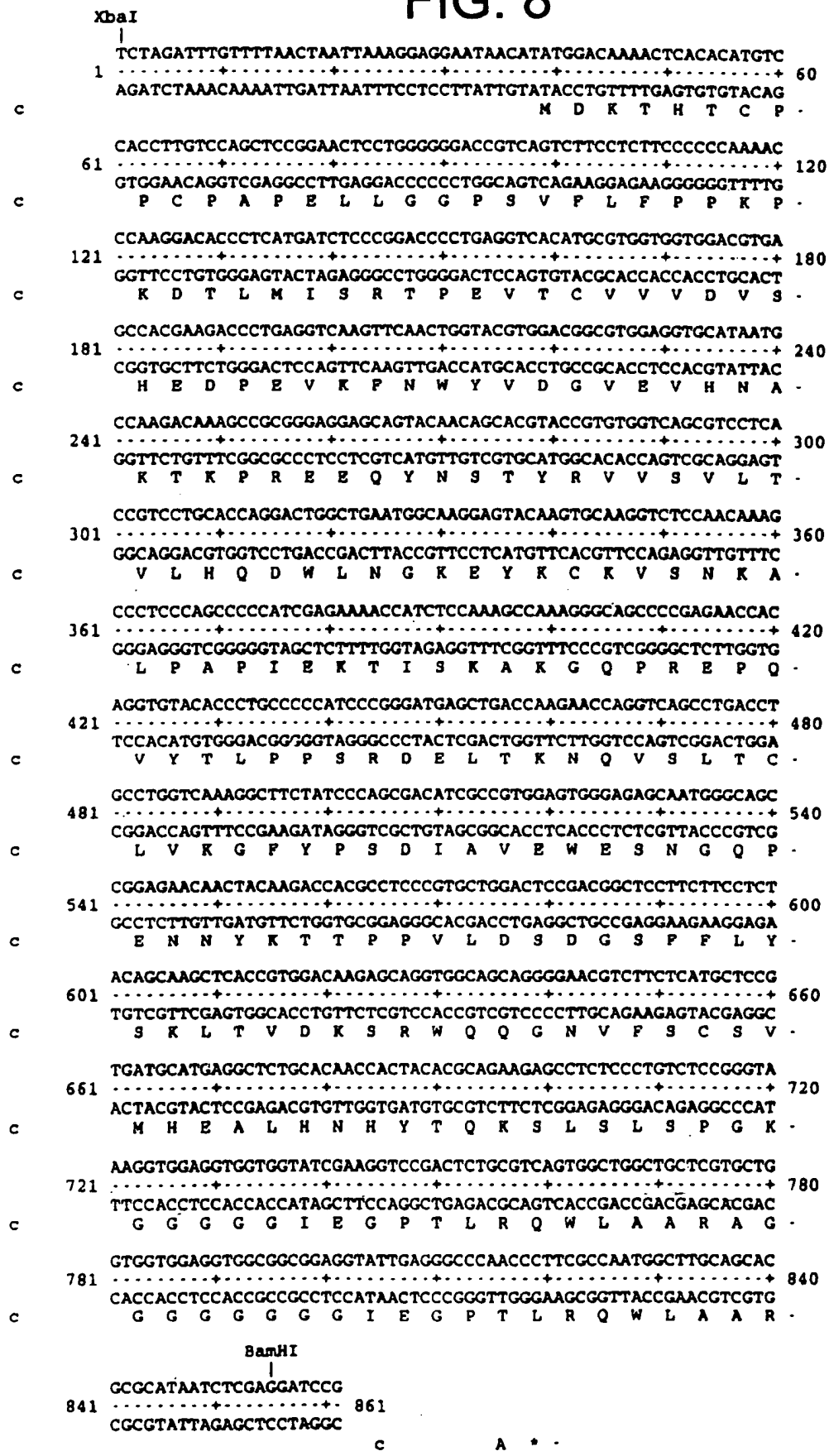


FIG. 9

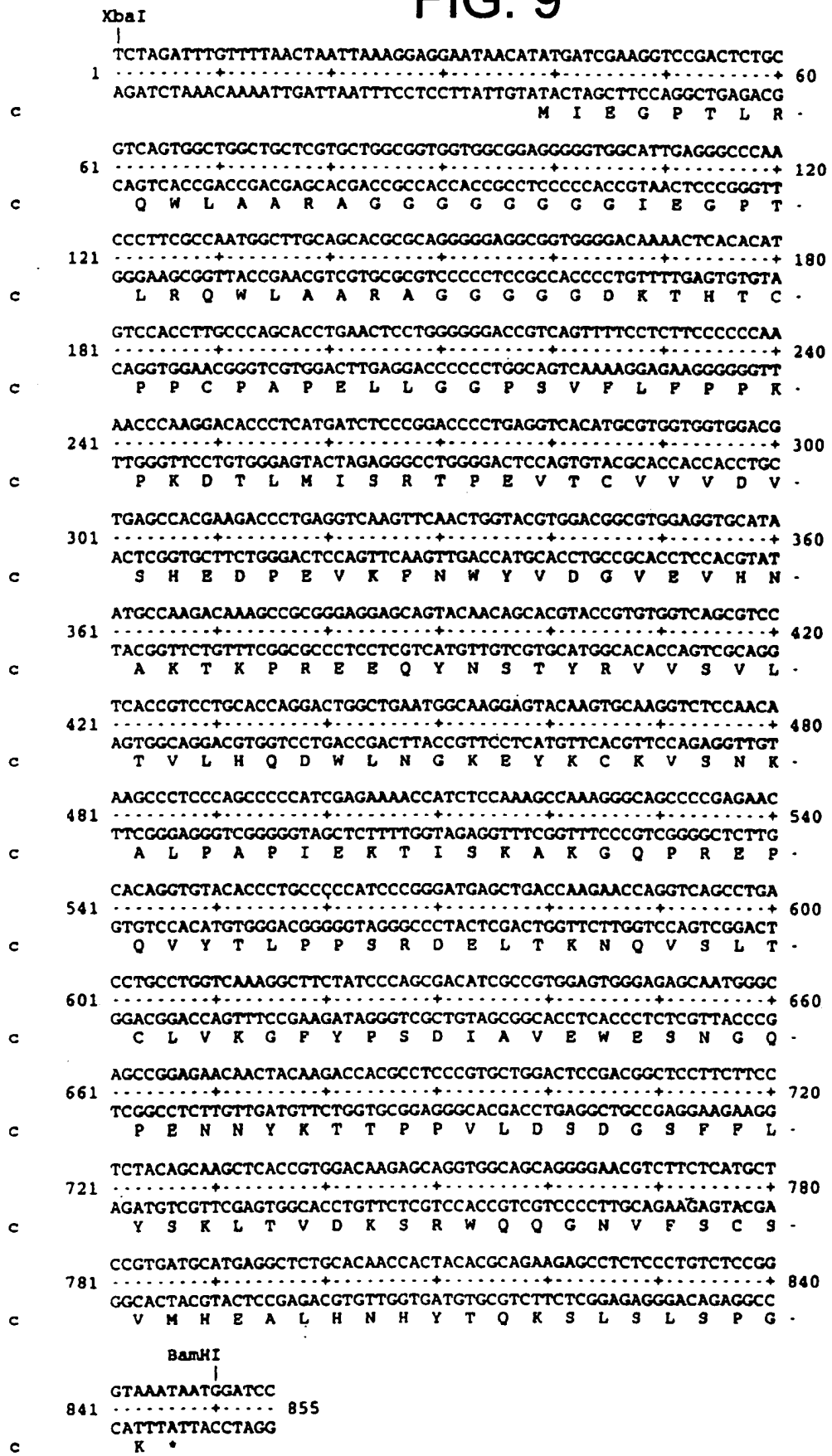

XbaI
TCTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGATCGAAGGTCCGACTCTGC
AGATCTAAACAAAATTGATTAATTTCCTCCTTATTGTATACTAGCTTCCAGGCTGAGACG
M I E G P T L R
GTCAGTGGCTGGCTGCTCGTGCTGGCGGTGGTGGCGGAGGGGGTGGCATTGAGGGCCCAA
CAGTCACCGACCGACGAGCACGACCGCCACCACCGCCTCCCCCACCGTAACTCCCGGGTT
Q W L A A R A G G G G G G G G I E G P T
CCCTTCGCCAATGGCTTGCAGCACGCGCAGGGGGAGGCGGTGGGGACAAAACTCACACAT
GGGAAGCGGTTACCGAACGTCGTGCGCGTCCCCCTCCGCCACCCCTGTTTTGAGTGTGTA
L R Q W L A A R A G G G G G D K T H T C
GTCCACCTTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTTTTCCTCTTCCCCCCAA
CAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGCAGTCAAAAGGAGAAGGGGGGTT
P P C P A P E L L G G P S V F L F P P P K
AACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACG
TTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGC
P K D T L M I S R T P E V T C V V V D V
TGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA
ACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTAT
S H E D P E V K F N W Y V D G V E V H N
ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCC
TACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCACACCAGTCGCAGG
A K T K P R E E Q Y N S T Y R V V S V L
TCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACA
AGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGT
T V L H Q D W L N G K E Y K C K V S N K
AAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
TTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTG
A L P A P I E K T I S K A K G Q P R E P
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGA
GTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACT
Q V Y T L P P S R D E L T K N Q V S L T
CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
GGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCG
C L V K G F Y P S D I A V E W E S N G Q
AGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCC
TCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGG
P E N N Y K T T P P V L D S D G S F F L
TCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCT
AGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGA
Y S K L T V D K S R W Q Q G N V F S C S
CCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
GGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCC
V M H E A L H N H Y T Q K S L S L S P G
BamHI
GTAAATAATGGATCC
CATTTATTACCTAGG
K *

FIG. 10

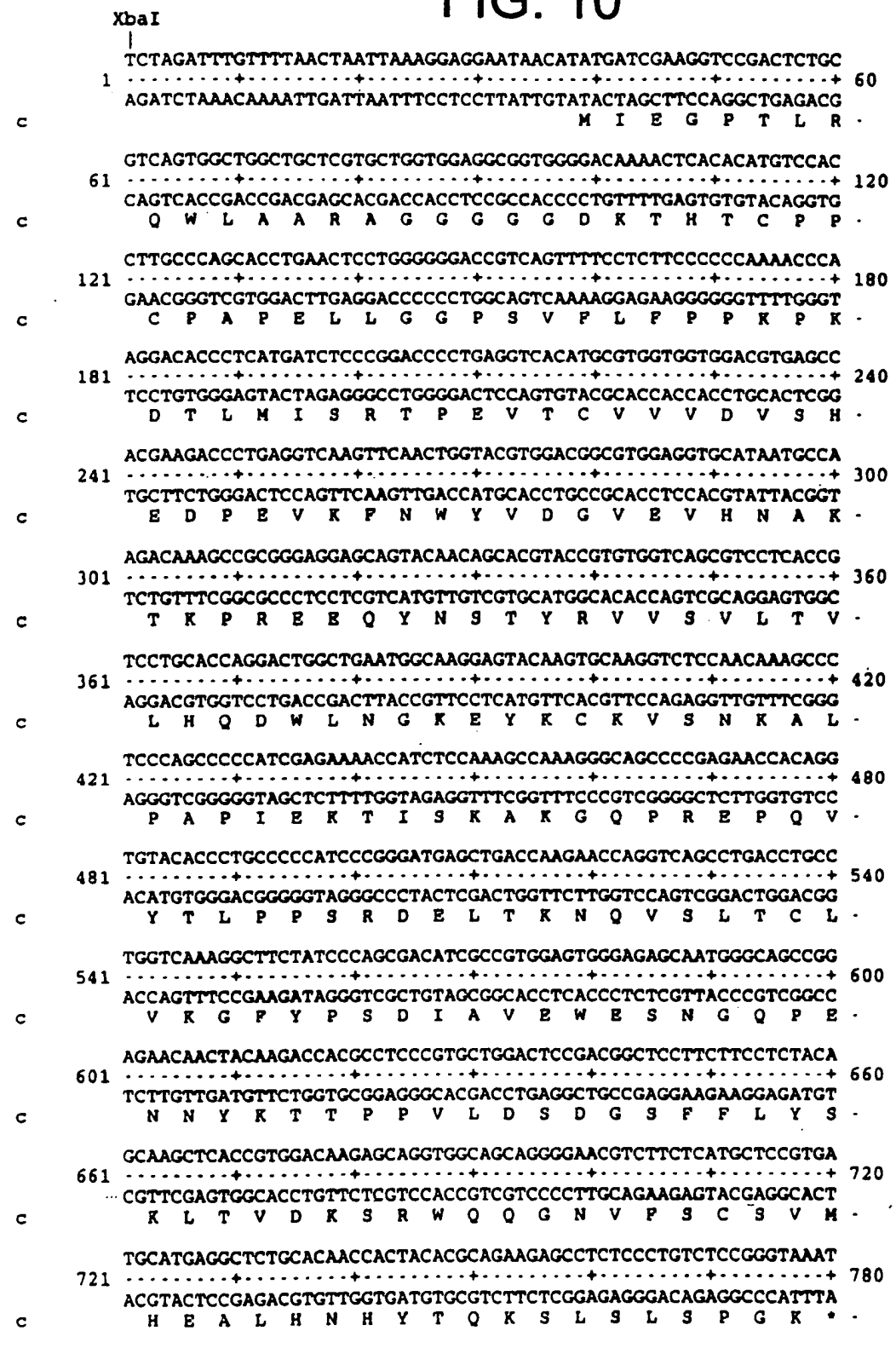


FIG.11

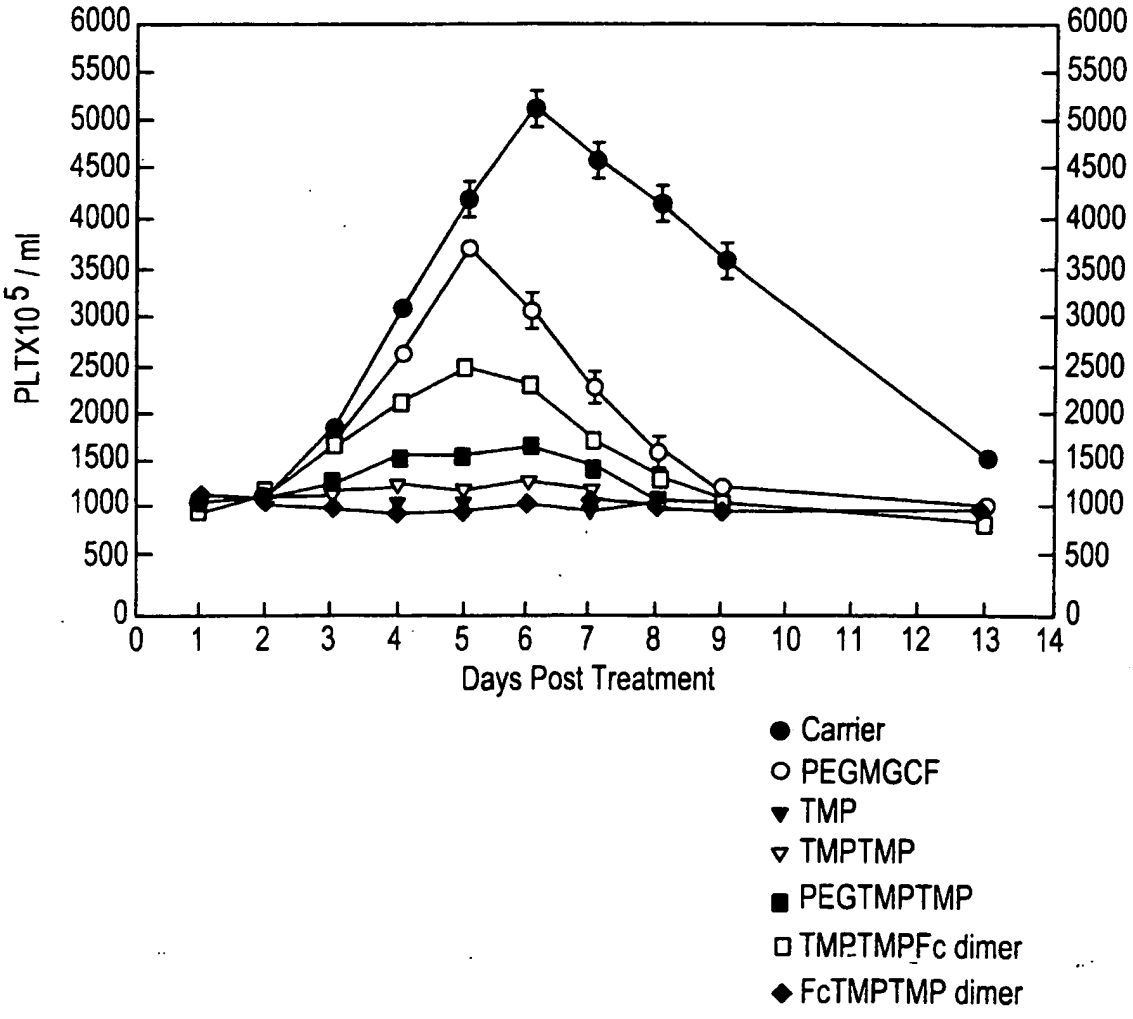

PLTX10^5/ml
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Days Post Treatment
Carrier
PEGMGCF
TMP
TMPTMP
PEGTMPTMP
TMPTMPFc dimer
FcTMPTMP dimer

FIG.12

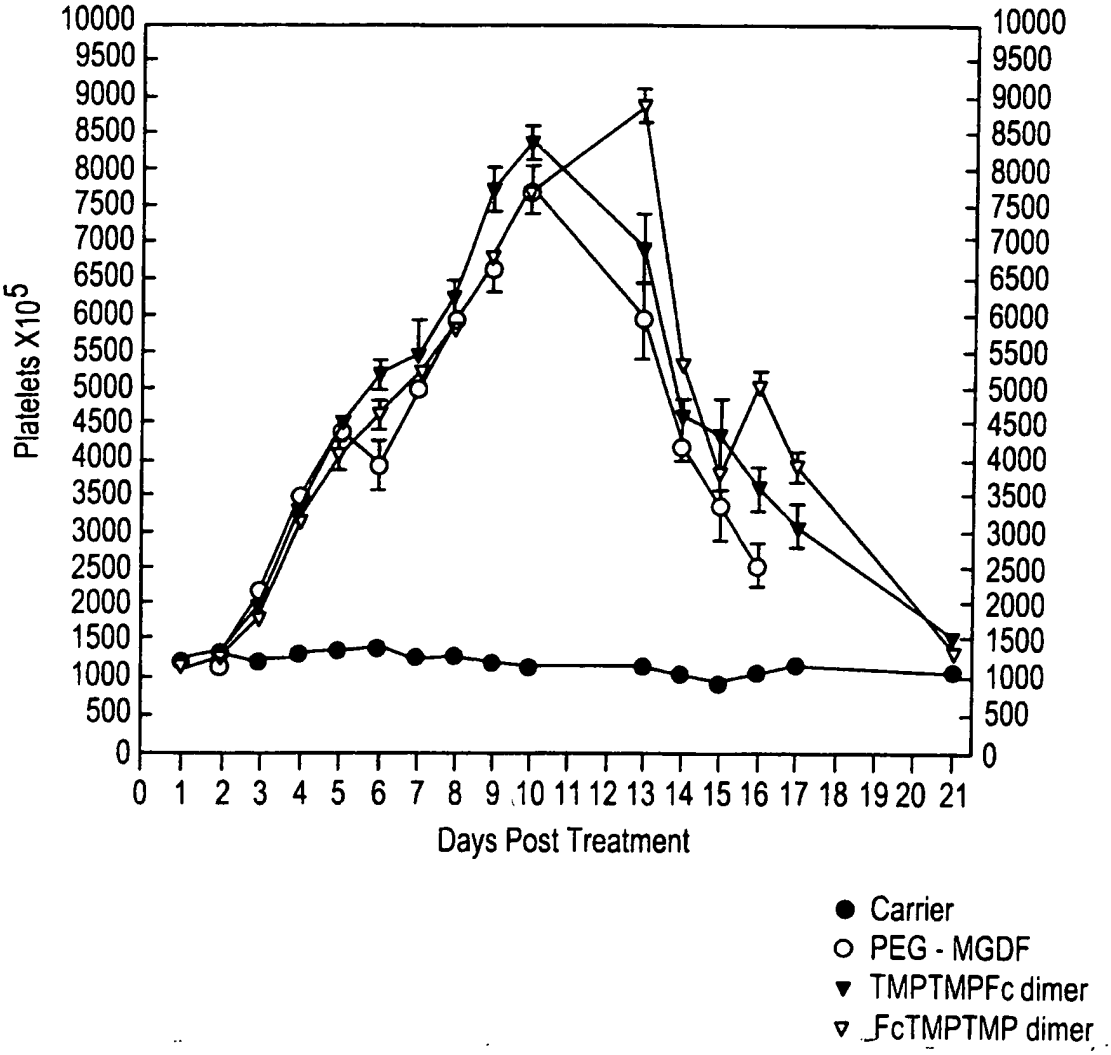

Platelets X10^5
10000
9500
9000
8500
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Days Post Treatment
Carrier
PEG - MGDF
TMPTMPFc dimer
FcTMPTMP dimer

FIG. 13

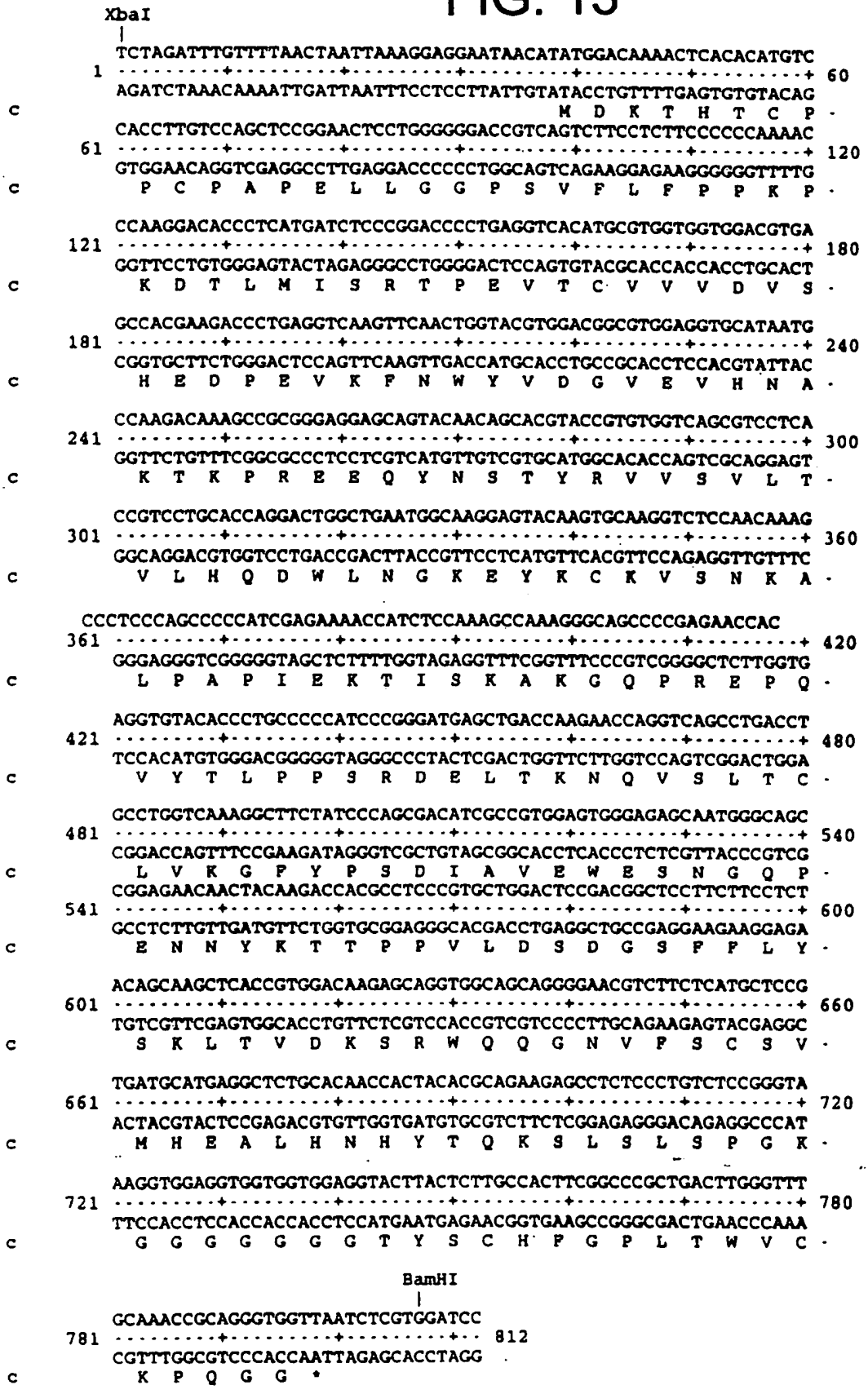

XbaI
TCTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGGACAAAACTCACACATGTC
AGATCTAAACAAAATTGATTAATTTCCTCCTTATTGTATACCTGTTTTGAGTGTGTACAG
M D K T H T C P
CACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAAC
GTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGCAGTCAGAAGGAGAAGGGGGGTTTTG
P C P A P E L L G G P S V F L F P P K P
CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGA
GGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACT
K D T L M I S R T P E V T C V V V D V S
GCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
CGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTAC
H E D P E V K F N W Y V D G V E V H N A
CCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA
GGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCACACCAGTCGCAGGAGT
K T K P R E E Q Y N S T Y R V V S V L T
CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAG
GGCAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTC
V L H Q D W L N G K E Y K C K V S N K A
CCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCAC
GGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTG
L P A P I E K T I S K A K G Q P R E P Q
AGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCT
TCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGA
V Y T L P P S R D E L T K N Q V S L T C
GCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGC
CGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCG
L V K G F Y P S D I A V E W E S N G Q P
CGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCT
GCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGA
E N N Y K T T P P V L D S D G S F F L Y
ACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG
TGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGC
S K L T V D K S R W Q Q G N V F S C S V
TGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA
ACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCAT
M H E A L H N H Y T Q K S L S L S P G K
AAGGTGGAGGTGGTGGTGGAGGTACTTACTCTTGCCACTTCGGCCCGCTGACTTGGGTTT
TTCCACCTCCACCACCACCTCCATGAATGAGAACGGTGAAGCCGGGCGACTGAACCCAAA
G G G G G G G T Y S C H F G P L T W V C
BamHI
GCAAACCGCAGGGTGGTTAATCTCGTGGATCC
CGTTTGGCGTCCCACCAATTAGAGCACCTAGG
K P Q G G *

FIG. 14

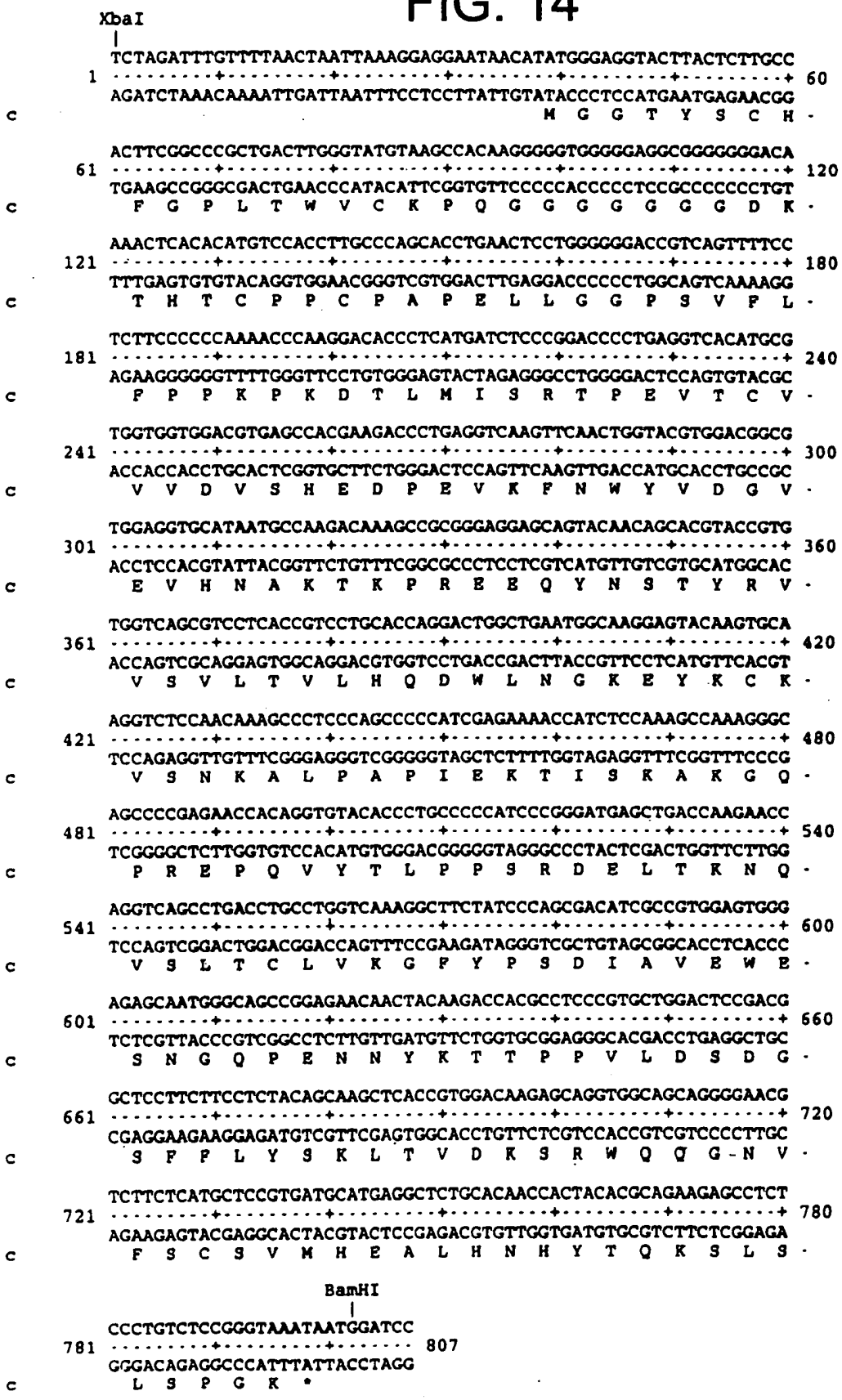

XbaI
TCTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGGGAGGTACTTACTCTTGCC
AGATCTAAACAAAATTGATTAATTTCCTCCTTATTGTATACCCTCCATGAATGAGAACGG
M G G T Y S C H
ACTTCGGCCCGCTGACTTGGGTATGTAAGCCACAAGGGGGTGGGGGAGGCGGGGGGGACA
TGAAGCCGGGCGACTGAACCCATACATTCGGTGTTCCCCCACCCCCTCCGCCCCCCCTGT
F G P L T W V C K P Q G G G G G G G D K
AAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTTTTCC
TTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGCAGTCAAAAGG
T H T C P P C P A P E L L G G P S V F L
TCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCG
AGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGC
F P P K P K D T L M I S R T P E V T C V
TGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCG
ACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGC
V V D V S H E D P E V K F N W Y V D G V
TGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTG
ACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCAC
E V H N A K T K P R E E Q Y N S T Y R V
TGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCA
ACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGT
V S V L T V L H Q D W L N G K E Y K C K
AGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGC
TCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCG
V S N K A L P A P I E K T I S K A K G Q
AGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACC
TCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGG
P R E P Q V Y T L P P S R D E L T K N Q
AGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG
TCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCC
V S L T C L V K G F Y P S D I A V E W E
AGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACG
TCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGC
S N G Q P E N N Y K T T P P V L D S D G
GCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACG
CGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGC
S F F L Y S K L T V D K S R W Q Q G N V
TCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCT
AGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGA
F S C S V M H E A L H N H Y T Q K S L S
BamHI
CCCTGTCTCCGGGTAAATAATGGATCC
GGGACAGAGGCCCATTTATTACCTAGG
L S P G K *

FIG. 15

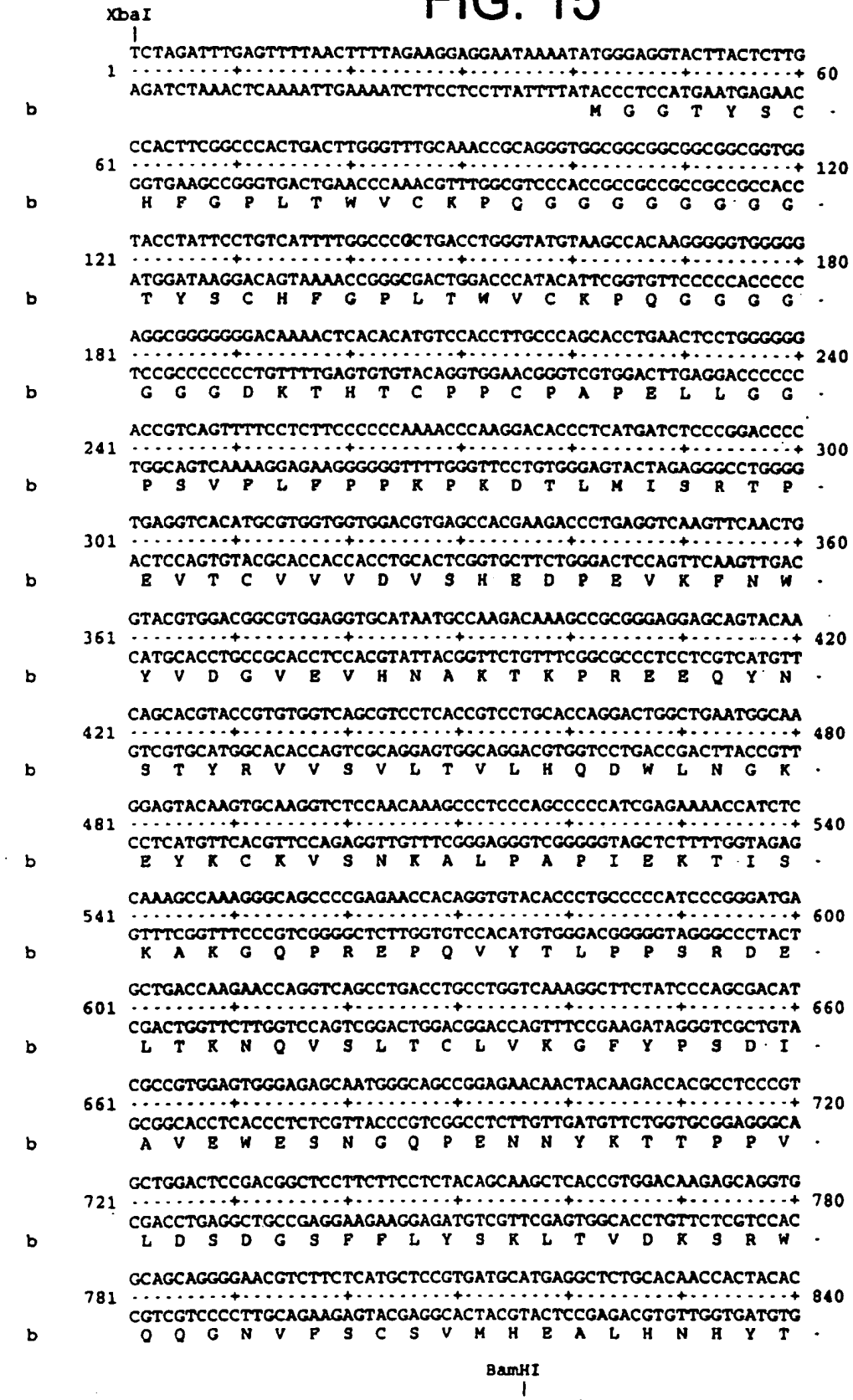


FIG. 16

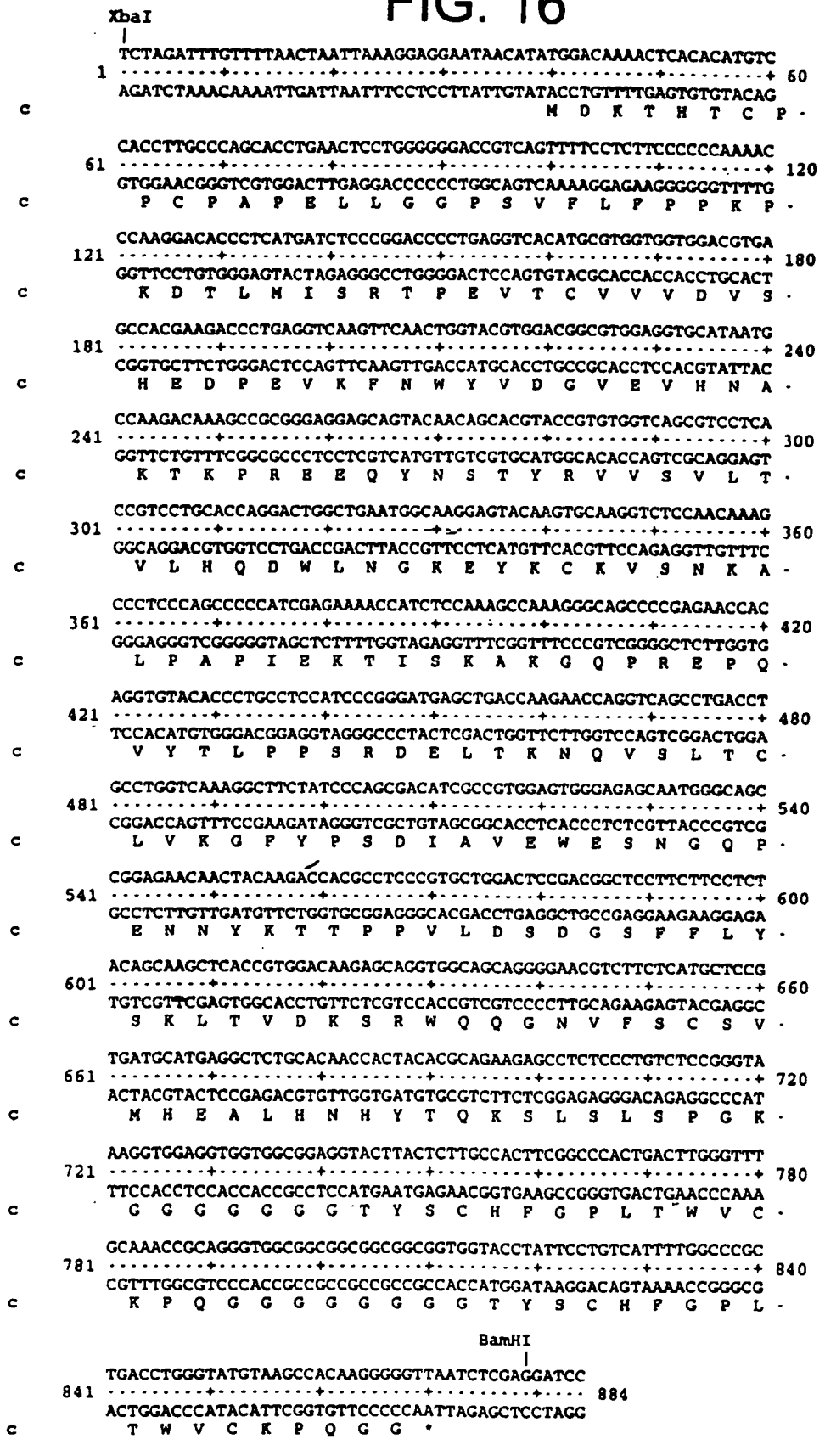


FIG. 17A

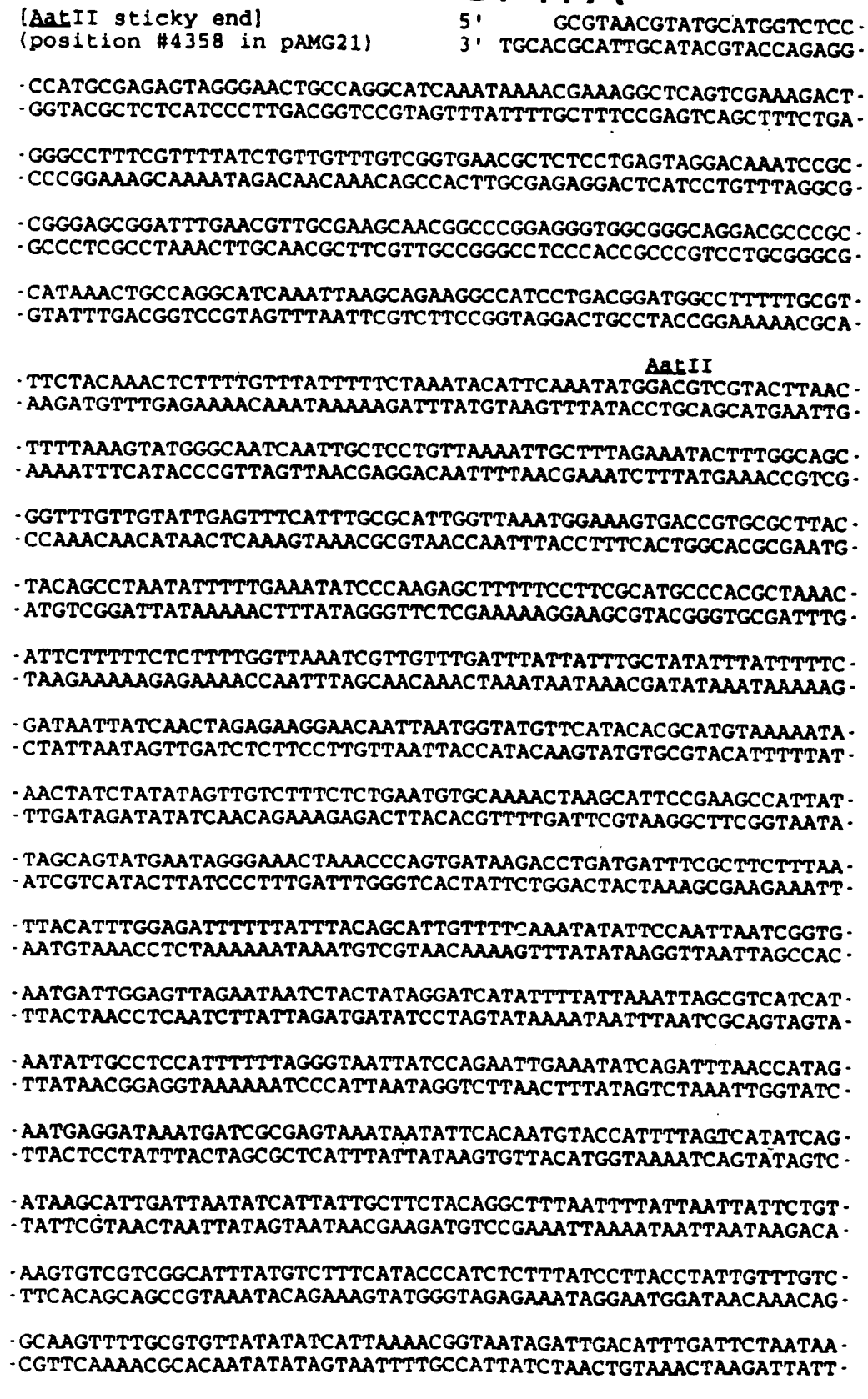

[AatII sticky end]
(position #4358 in pAMG21)
5' GCGTAACGTATGCATGGTCTCC
3' TGCACGCATTGCATACGTACCAGAGG
AatII

FIG. 17B

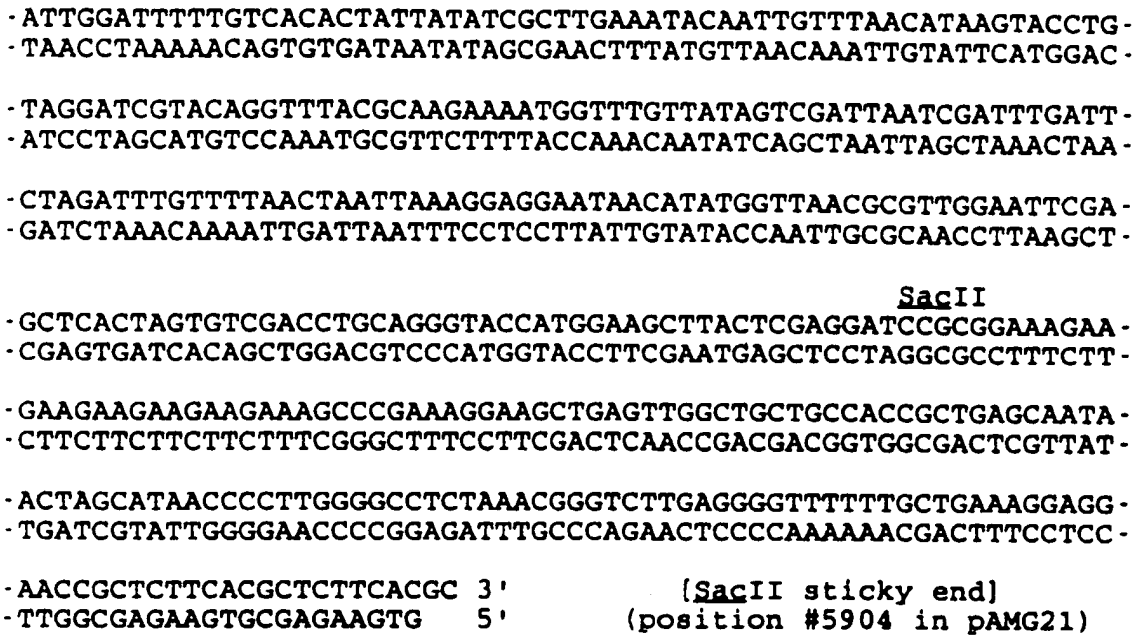

-ATTGGATTTTTGTCACACTATTATATCGCTTGAAATACAATTGTTTAACATAAGTACCTG-
-TAACCTAAAAACAGTGTGATAATATAGCGAACTTTATGTTAACAAATTGTATTCATGGAC-
-TAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGATTAATCGATTTGATT-
-ATCCTAGCATGTCCAAATGCGTTCTTTTACCAAACAATATCAGCTAATTAGCTAAACTAA-
-CTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGGTTAACGCGTTGGAATTCGA-
-GATCTAAACAAAATTGATTAATTTCCTCCTTATTGTATACCAATTGCGCAACCTTAAGCT-
SacII
-GCTCACTAGTGTCGACCTGCAGGGTACCATGGAAGCTTACTCGAGGATCCGCGGAAAGAA-
-CGAGTGATCACAGCTGGACGTCCCATGGTACCTTCGAATGAGCTCCTAGGCGCCTTTCTT-
-GAAGAAGAAGAAGAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATA-
-CTTCTTCTTCTTCTTTCGGGCTTTCCTTCGACTCAACCGACGACGGTGGCGACTCGTTAT-
-ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGG-
-TGATCGTATTGGGGAACCCCGGAGATTTGCCCAGAACTCCCCAAAAAACGACTTTCCTCC-
-AACCGCTCTTCACGCTCTTCACGC 3'
-TTGGCGAGAAGTGCGAGAAGTG 5'
[SacII sticky end]
(position #5904 in pAMG21)

FIG.18A - 1

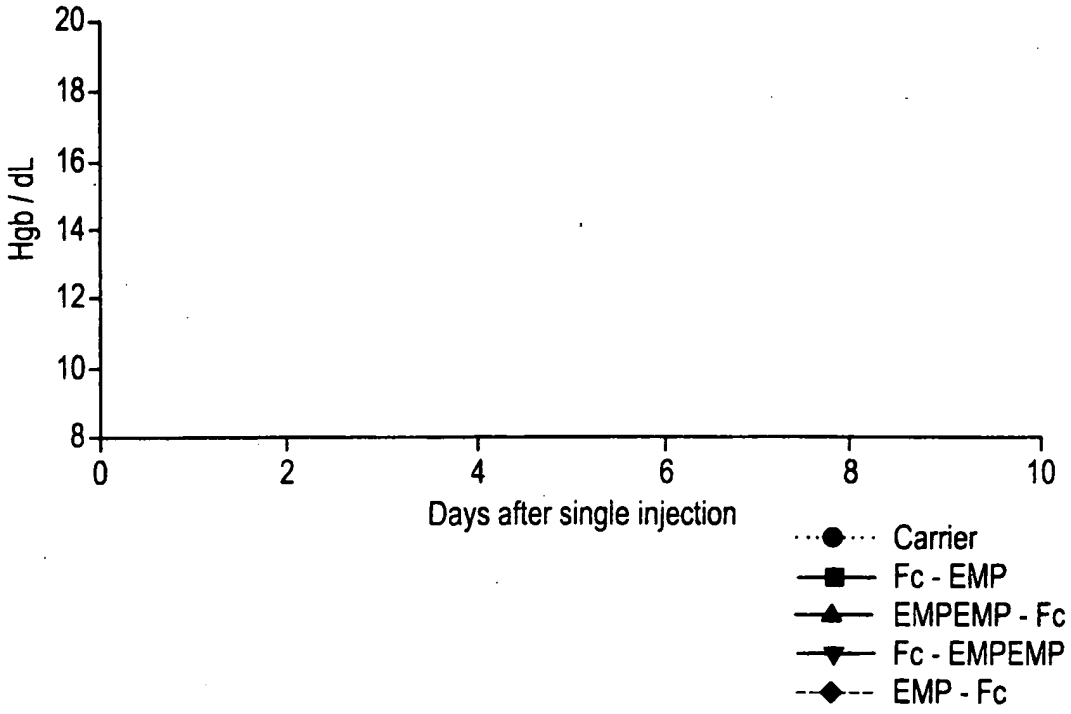


FIG.18A - 2

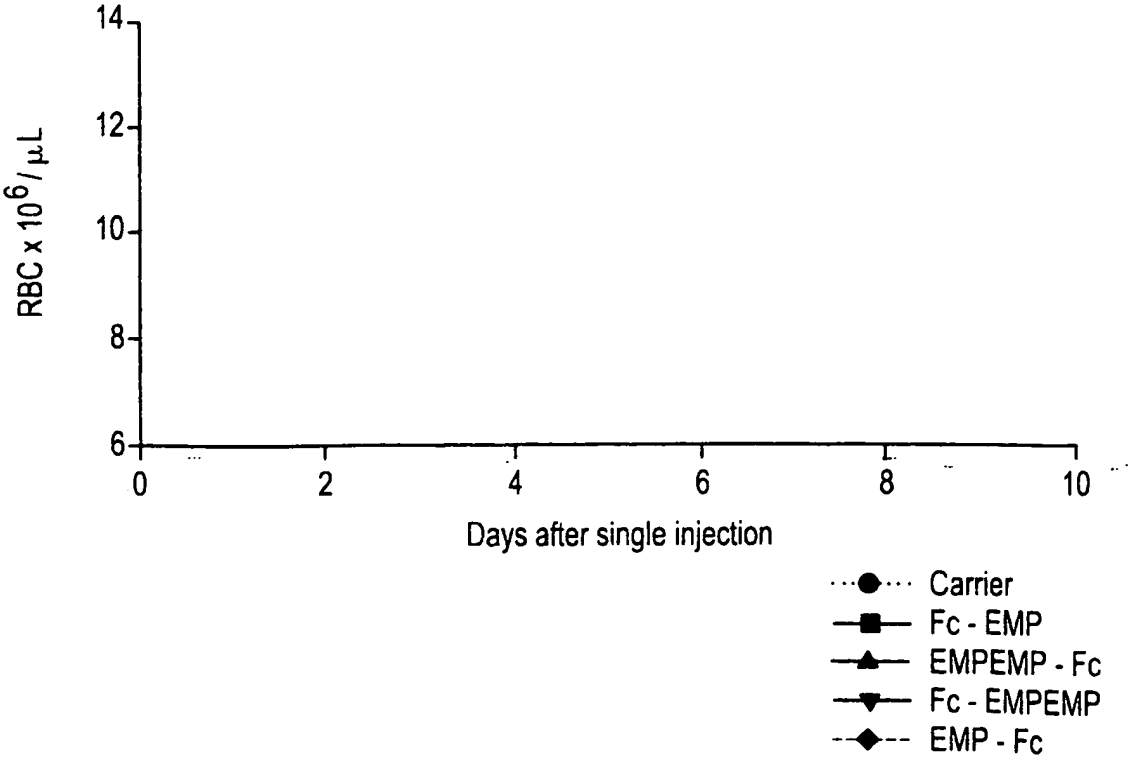


FIG.18A - 3

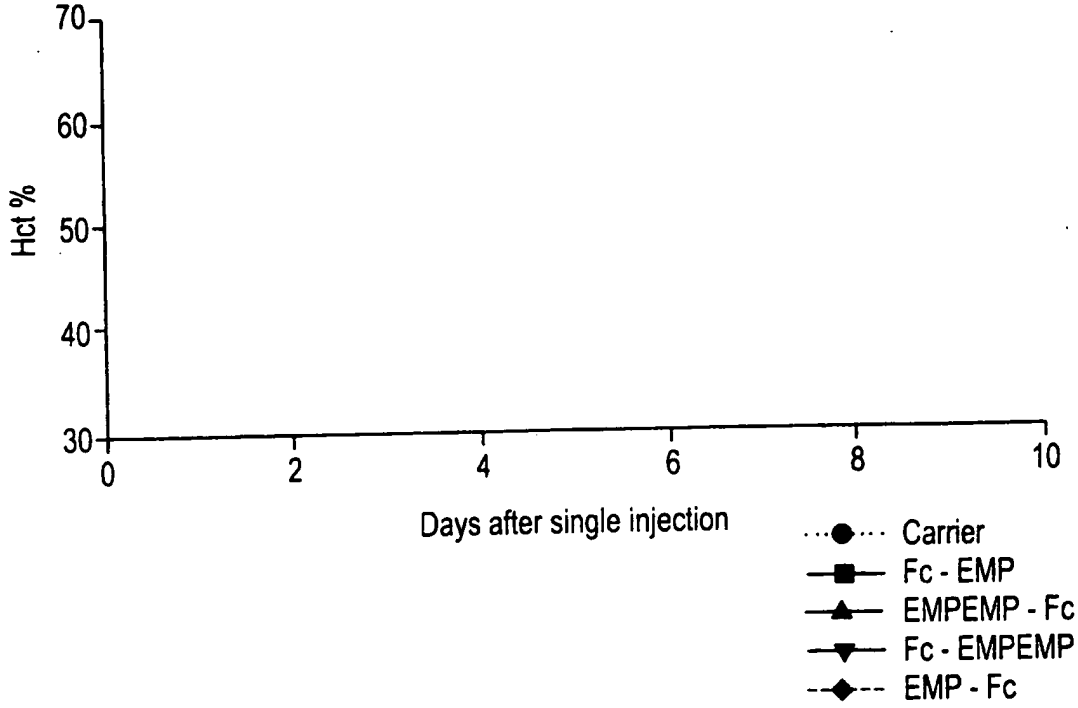


FIG.18B - 1

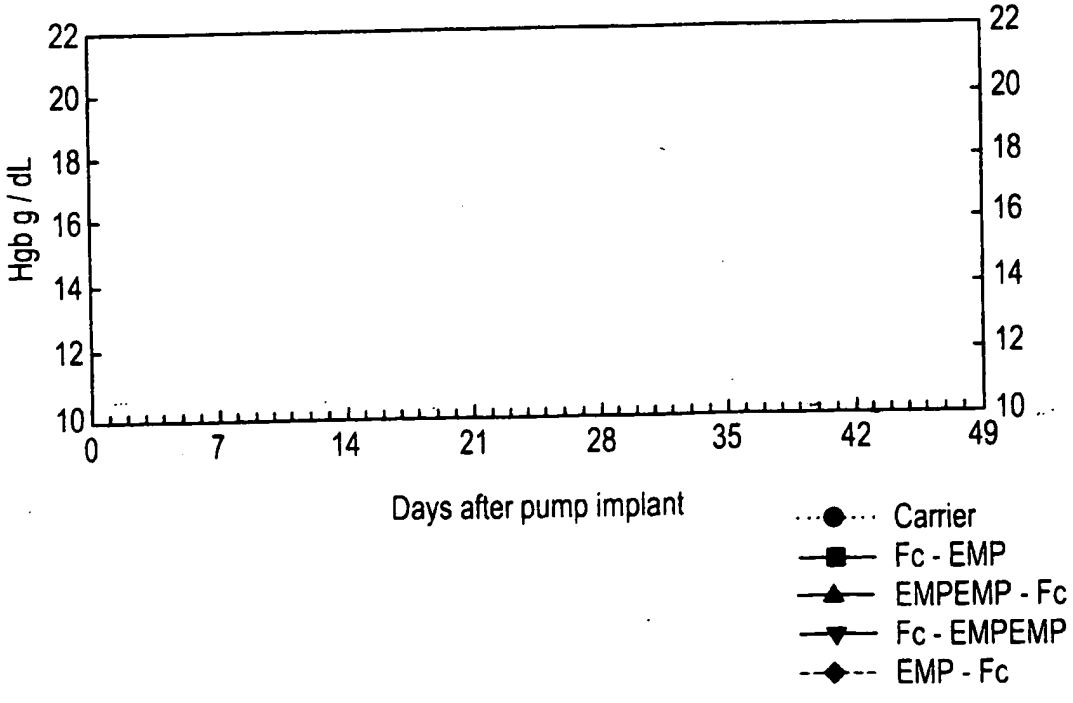


FIG.18B - 2

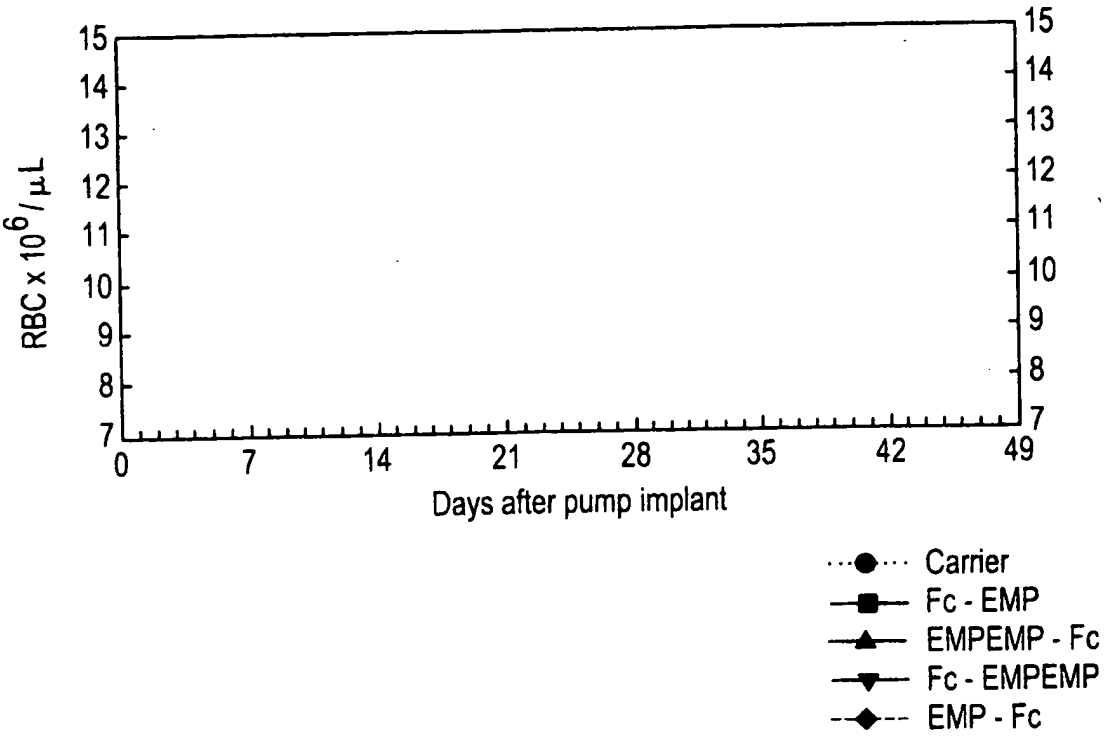


FIG.18B - 3

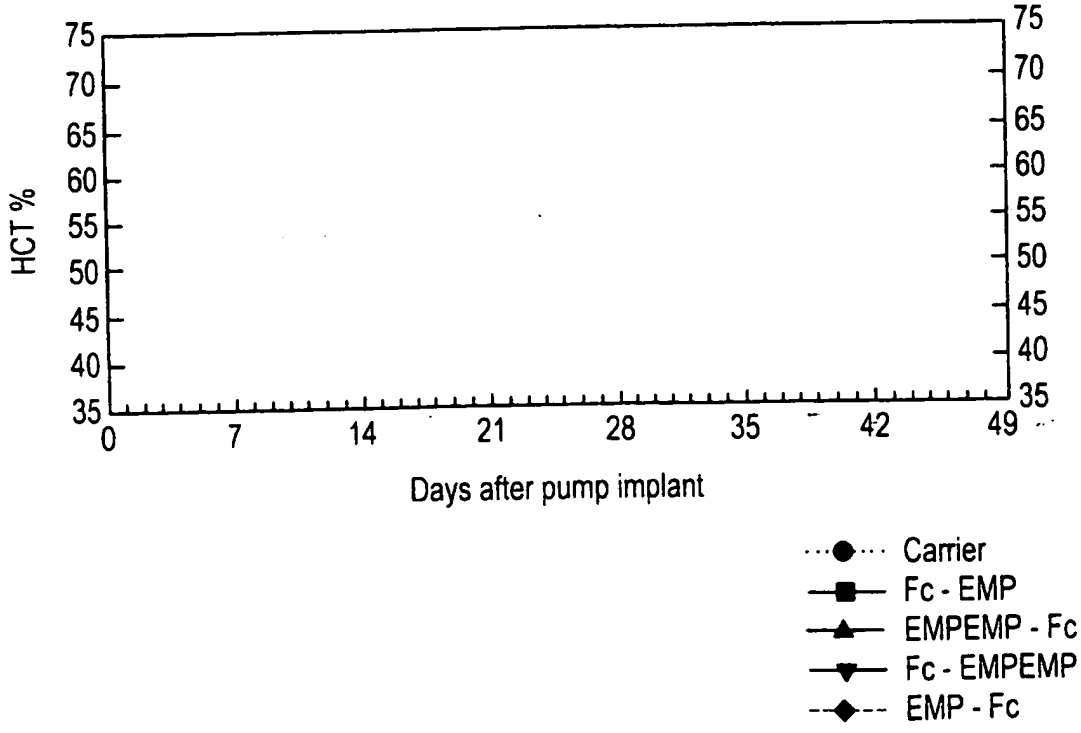


FIG. 19A

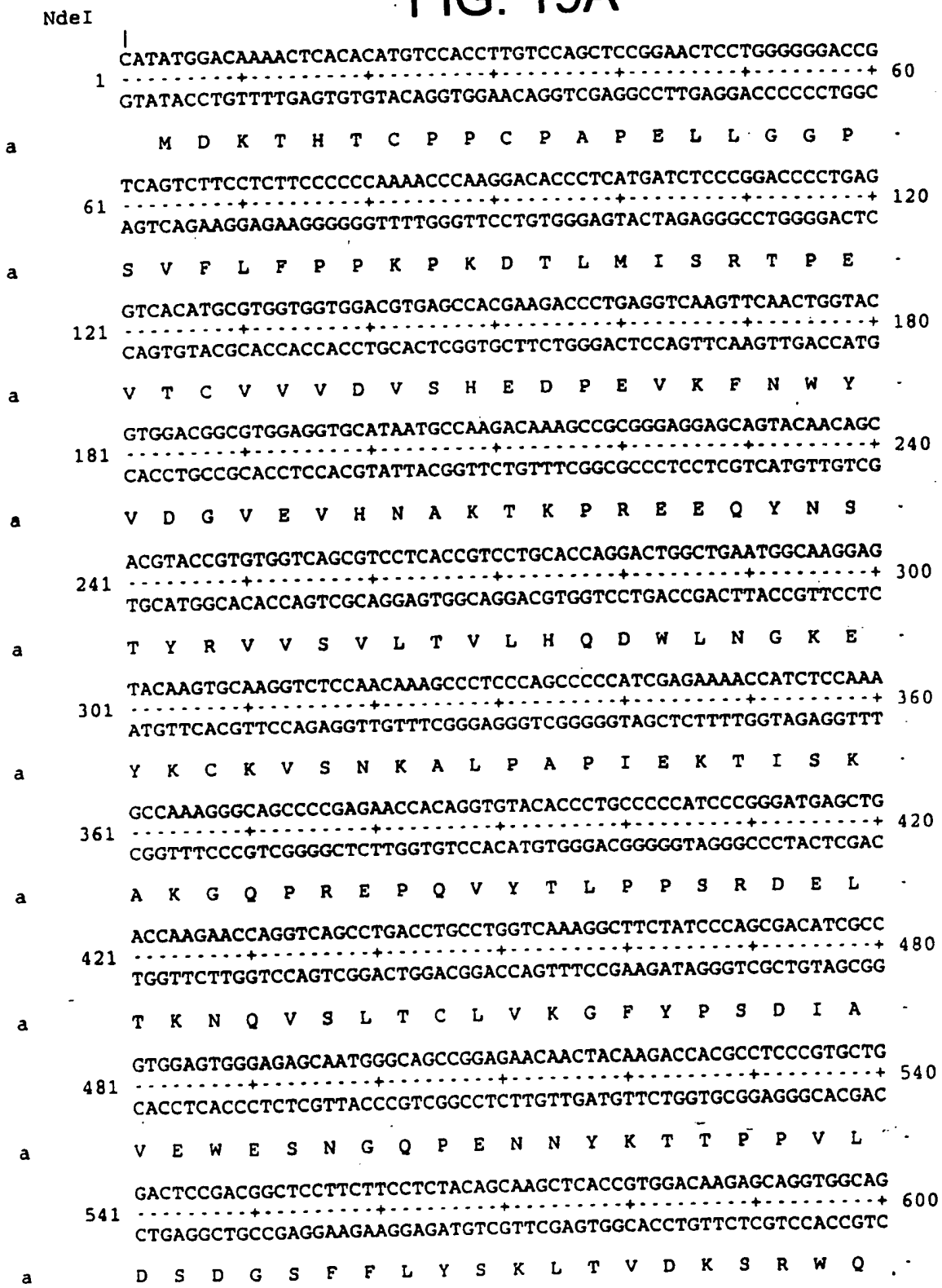

NdeI
CATATGGACAAAACTCACACATGTCCACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCG
GTATACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGC
M D K T H T C P P C P A P E L L G G P
TCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
AGTCAGAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
S V F L F P P K P K D T L M I S R T P E
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
V T C V V V D V S H E D P E V K F N W Y
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
V D G V E V H N A K T K P R E E Q Y N S
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
T Y R V V S V L T V L H Q D W L N G K E
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
Y K C K V S N K A L P A P I E K T I S K
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
A K G Q P R E P Q V Y T L P P S R D E L
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
T K N Q V S L T C L V K G F Y P S D I A
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
V E W E S N G Q P E N N Y K T T P P V L
GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
D S D G S F F L Y S K L T V D K S R W Q

FIG. 19B

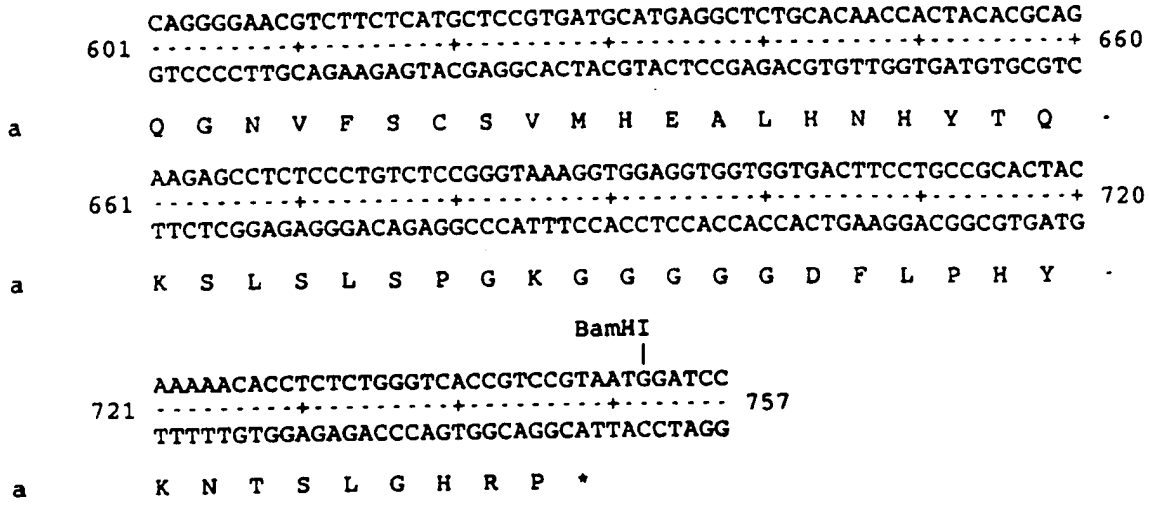

CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
601
660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a Q G N V F S C S V M H E A L H N H Y T Q
AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTGACTTCCTGCCGCACTAC
661
720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCACTGAAGGACGGCGTGATG
a K S L S L S P G K G G G G G D F L P H Y
BamHI
AAAAACACCTCTCTGGGTCACCGTCCGTAATGGATCC
721
757
TTTTTGTGGAGAGACCCAGTGGCAGGCATTACCTAGG
a K N T S L G H R P *

FIG. 20A

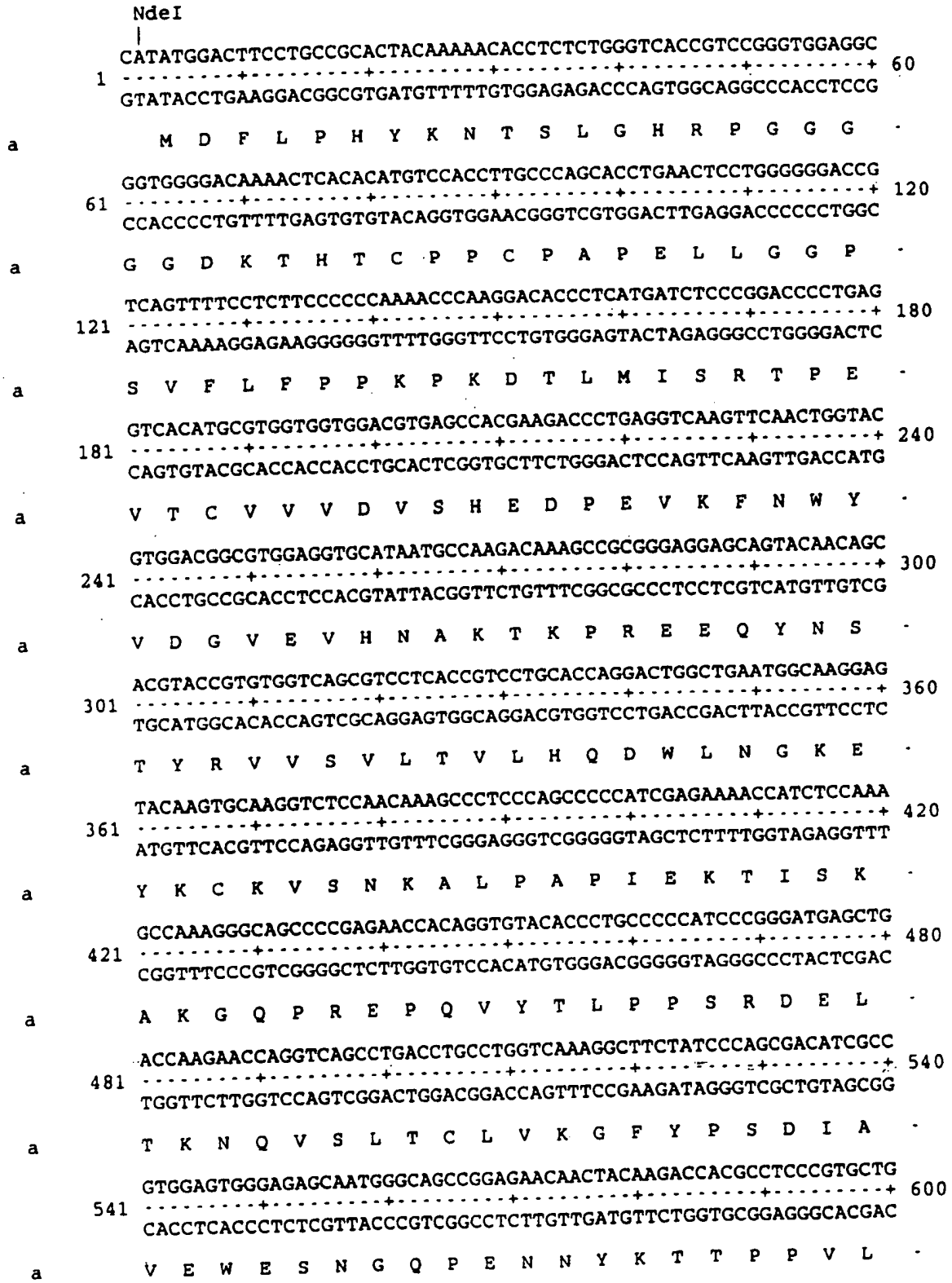

NdeI
CATATGGACTTCCTGCCGCACTACAAAAACACCTCTCTGGGTCACCGTCCGGGTGGAGGC
GTATACCTGAAGGACGGCGTGATGTTTTTGTGGAGAGACCCAGTGGCAGGCCCACCTCCG
M D F L P H Y K N T S L G H R P G G G
GGTGGGGACAAAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGGGGACCG
CCACCCCTGTTTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGC
G G D K T H T C P P C P A P E L L G G P
TCAGTTTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
AGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
S V F L F P P K P K D T L M I S R T P E
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
V T C V V V D V S H E D P E V K F N W Y
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
V D G V E V H N A K T K P R E E Q Y N S
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
T Y R V V S V L T V L H Q D W L N G K E
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
Y K C K V S N K A L P A P I E K T I S K
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
A K G Q P R E P Q V Y T L P P S R D E L
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
T K N Q V S L T C L V K G F Y P S D I A
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
V E W E S N G Q P E N N Y K T T P P V L
1
61
121
181
241
301
361
421
481
541
a
a
a
a
a
a
a
a
a

FIG. 20B

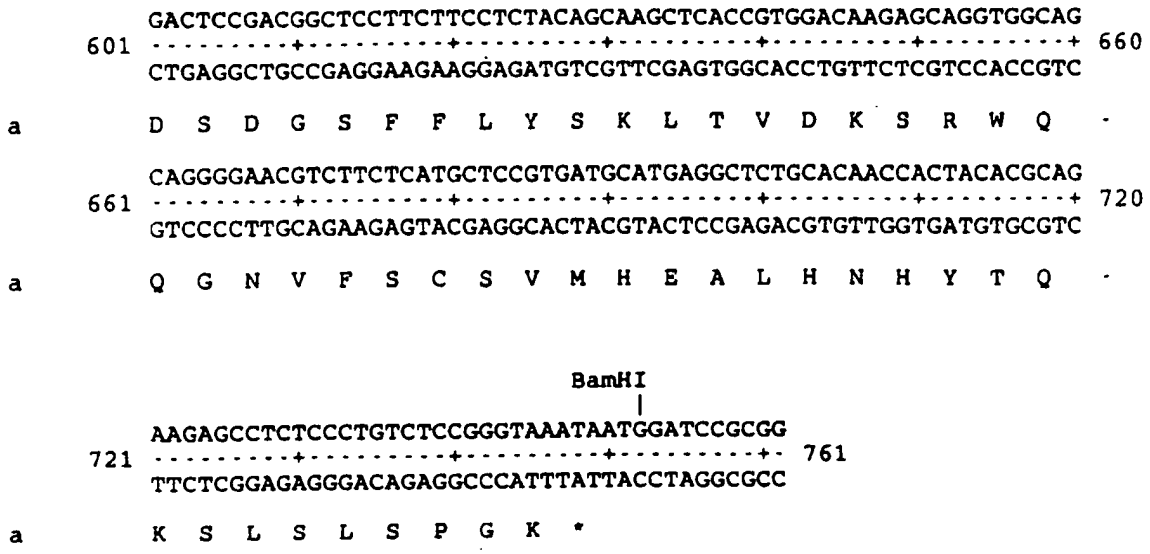

GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
601 660
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
a D S D G S F F L Y S K L T V D K S R W Q -
CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
661 720
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a Q G N V F S C S V M H E A L H N H Y T Q -
BamHI
AAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCCGCGG
721 761
TTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGGCGCC
a K S L S L S P G K *

FIG. 21A

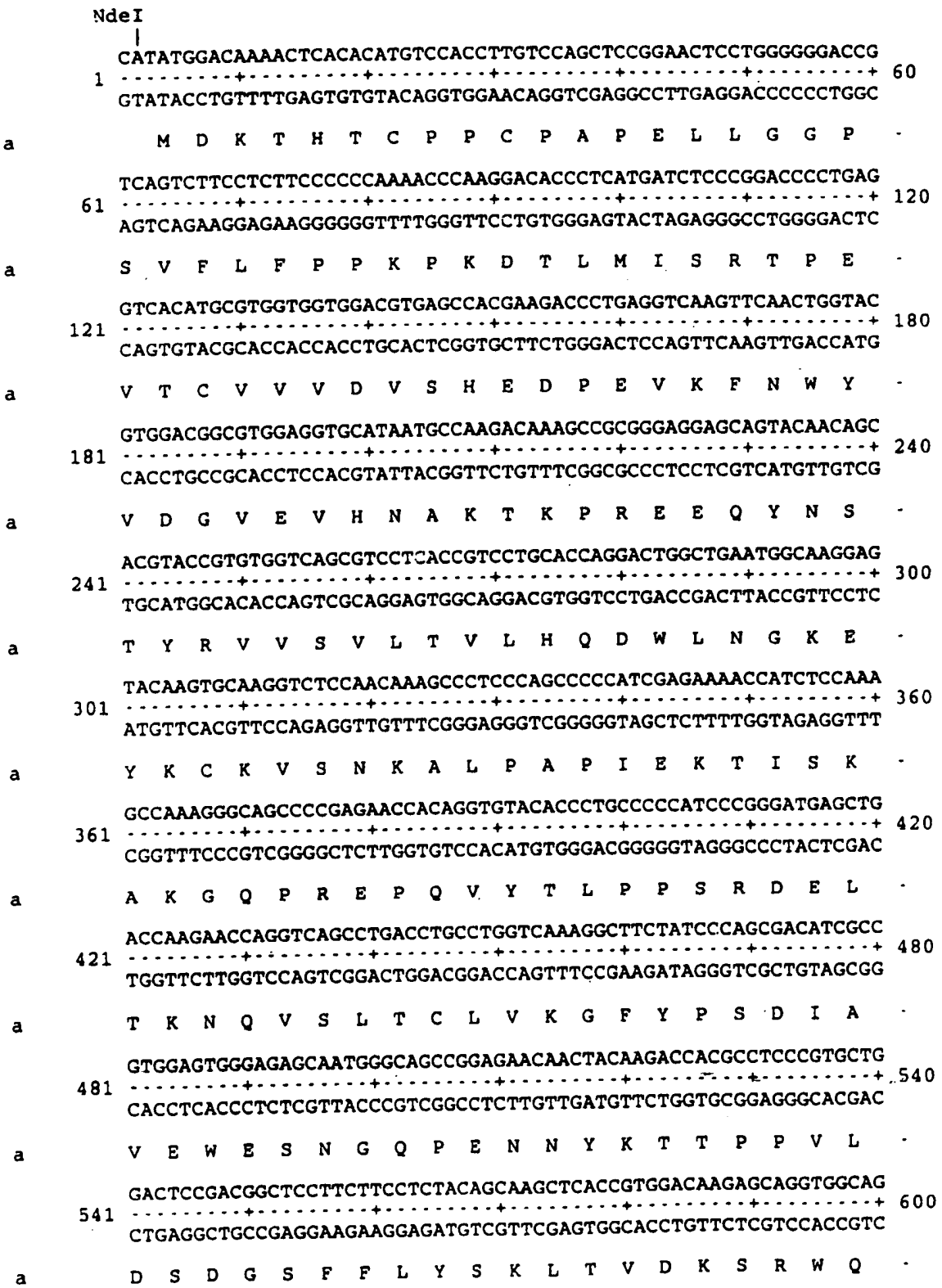

NdeI
CATATGGACAAAACTCACACATGTCCACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCG
GTATACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGC
a M D K T H T C P P C P A P E L L G G P -
TCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
AGTCAGAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
a S V F L F P P K P K D T L M I S R T P E -
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
a V T C V V V D V S H E D P E V K F N W Y -
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
a V D G V E V H N A K T K P R E E Q Y N S -
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
a T Y R V V S V L T V L H Q D W L N G K E -
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
a Y K C K V S N K A L P A P I E K T I S K -
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
CGGTTTCCCGTCGGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
a A K G Q P R E P Q V Y T L P P S R D E L -
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
a T K N Q V S L T C L V K G F Y P S D I A -
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
a V E W E S N G Q P E N N Y K T T P P V L -
GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
a D S D G S F F L Y S K L T V D K S R W Q -

FIG. 21B

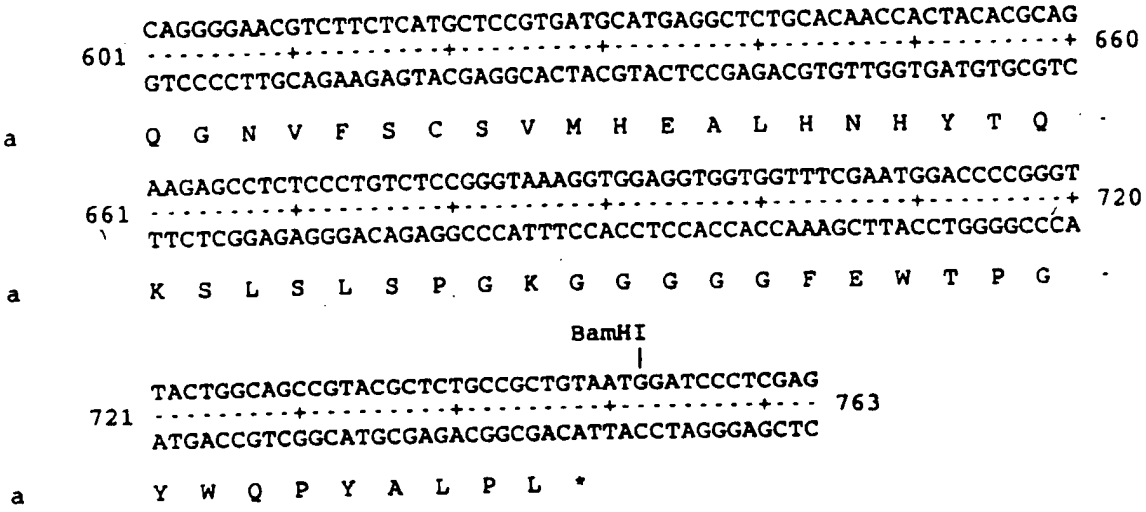

CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
601
660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a Q G N V F S C S V M H E A L H N H Y T Q
AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTTTCGAATGGACCCCGGGT
661
720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCAAAGCTTACCTGGGGCCCA
a K S L S L S P G K G G G G G F E W T P G
BamHI
TACTGGCAGCCGTACGCTCTGCCGCTGTAATGGATCCCTCGAG
721
763
ATGACCGTCGGCATGCGAGACGGCGACATTACCTAGGGAGCTC
a Y W Q P Y A L P L *

FIG. 22A

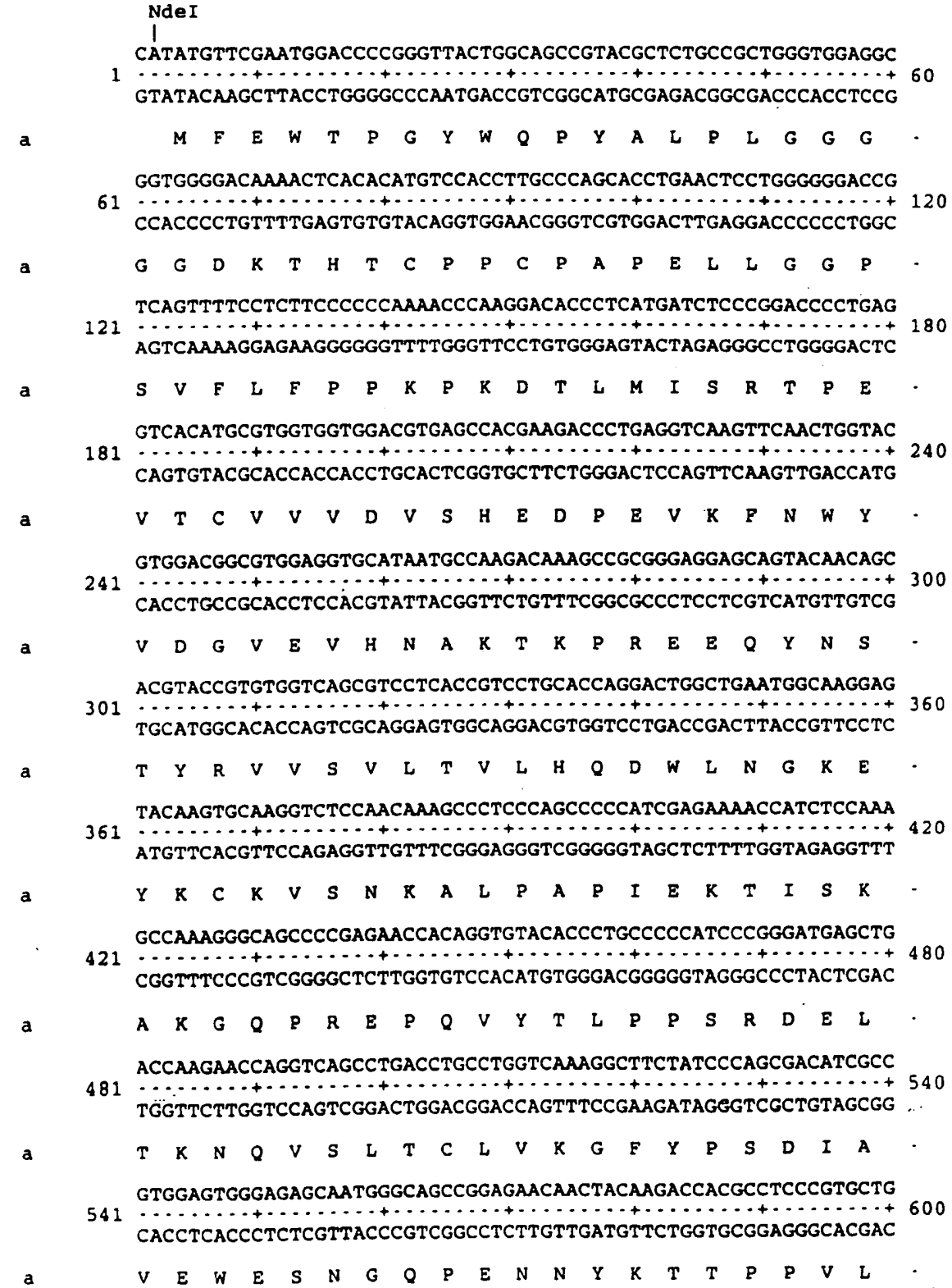

NdeI
CATATGTTCGAATGGACCCCGGGTTACTGGCAGCCGTACGCTCTGCCGCTGGGTGGAGGC
1
60
GTATACAAGCTTACCTGGGGCCCAATGACCGTCGGCATGCGAGACGGCGACCCACCTCCG
a M F E W T P G Y W Q P Y A L P L G G G
GGTGGGGACAAAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGGGGACCG
61
120
CCACCCCTGTTTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGC
a G G D K T H T C P P C P A P E L L G G P
TCAGTTTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
121
180
AGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
a S V F L F P P K P K D T L M I S R T P E
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
181
240
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
a V T C V V V D V S H E D P E V K F N W Y
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
241
300
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
a V D G V E V H N A K T K P R E E Q Y N S
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
301
360
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
a T Y R V V S V L T V L H Q D W L N G K E
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
361
420
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
a Y K C K V S N K A L P A P I E K T I S K
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
421
480
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
a A K G Q P R E P Q V Y T L P P S R D E L
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
481
540
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
a T K N Q V S L T C L V K G F Y P S D I A
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
541
600
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
a V E W E S N G Q P E N N Y K T T P P V L

FIG. 22B

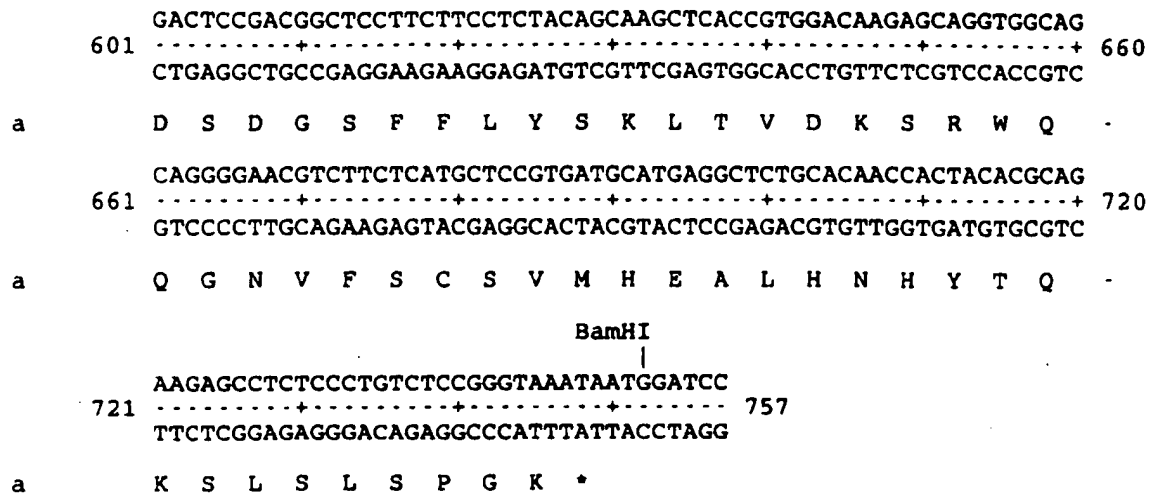

GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
601
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
660
a D S D G S F F L Y S K L T V D K S R W Q -
CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
661
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
720
a Q G N V F S C S V M H E A L H N H Y T Q -
BamHI
AAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCC
721
TTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGG
757
a K S L S L S P G K *

FIG. 23A

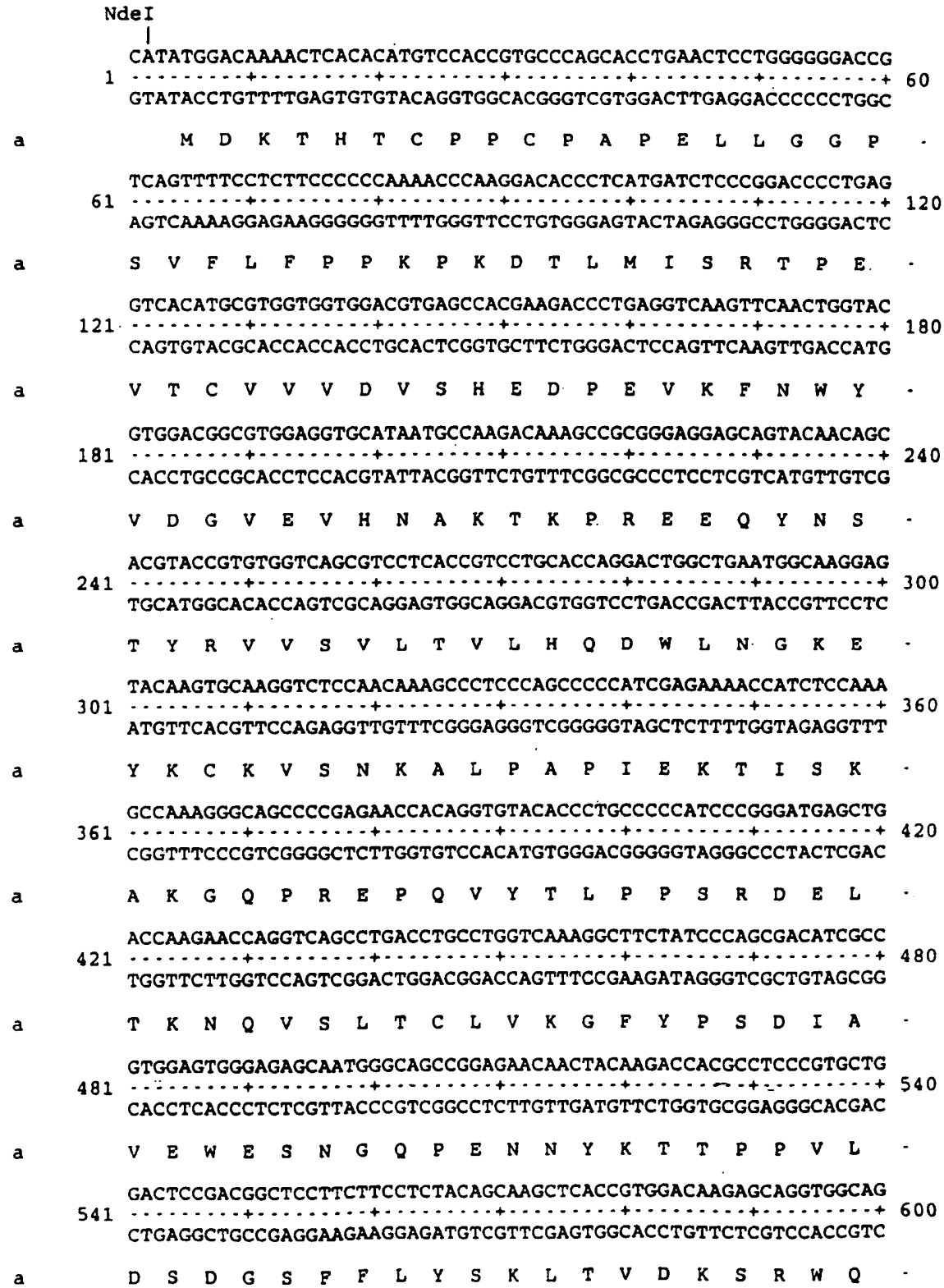

NdeI
CATATGGACAAAACTCACACATGTCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCG
GTATACCTGTTTTGAGTGTGTACAGGTGGCACGGGTCGTGGACTTGAGGACCCCCCTGGC
a M D K T H T C P P C P A P E L L G G P .
TCAGTTTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
AGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
a S V F L F P P K P K D T L M I S R T P E .
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
a V T C V V V D V S H E D P E V K F N W Y .
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
a V D G V E V H N A K T K P R E E Q Y N S .
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
a T Y R V V S V L T V L H Q D W L N G K E .
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
a Y K C K V S N K A L P A P I E K T I S K .
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
a A K G Q P R E P Q V Y T L P P S R D E L .
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
a T K N Q V S L T C L V K G F Y P S D I A .
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
a V E W E S N G Q P E N N Y K T T P P V L .
GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
a D S D G S F F L Y S K L T V D K S R W Q .
1
61
121
181
241
301
361
421
481
541
60
120
180
240
300
360
420
480
540
600

FIG. 23B

CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
601
660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a Q G N V F S C S V M H E A L H N H Y T Q
AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGTGGTGGTGGTGTTGAACCGAACTGTGAC
661
720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCACCACCACCACAACTTGGCTTGACACTG
a K S L S L S P G K G G G G G V E P N C D
BamHI
ATCCATGTTATGTGGGAATGGGAATGTTTTGAACGTCTGTAACTCGAGGATCC
721
773
TAGGTACAATACACCCTTACCCTTACAAAACTTGCAGACATTGAGCTCCTAGG
a I H V M W E W E C F E R L *

FIG. 24A

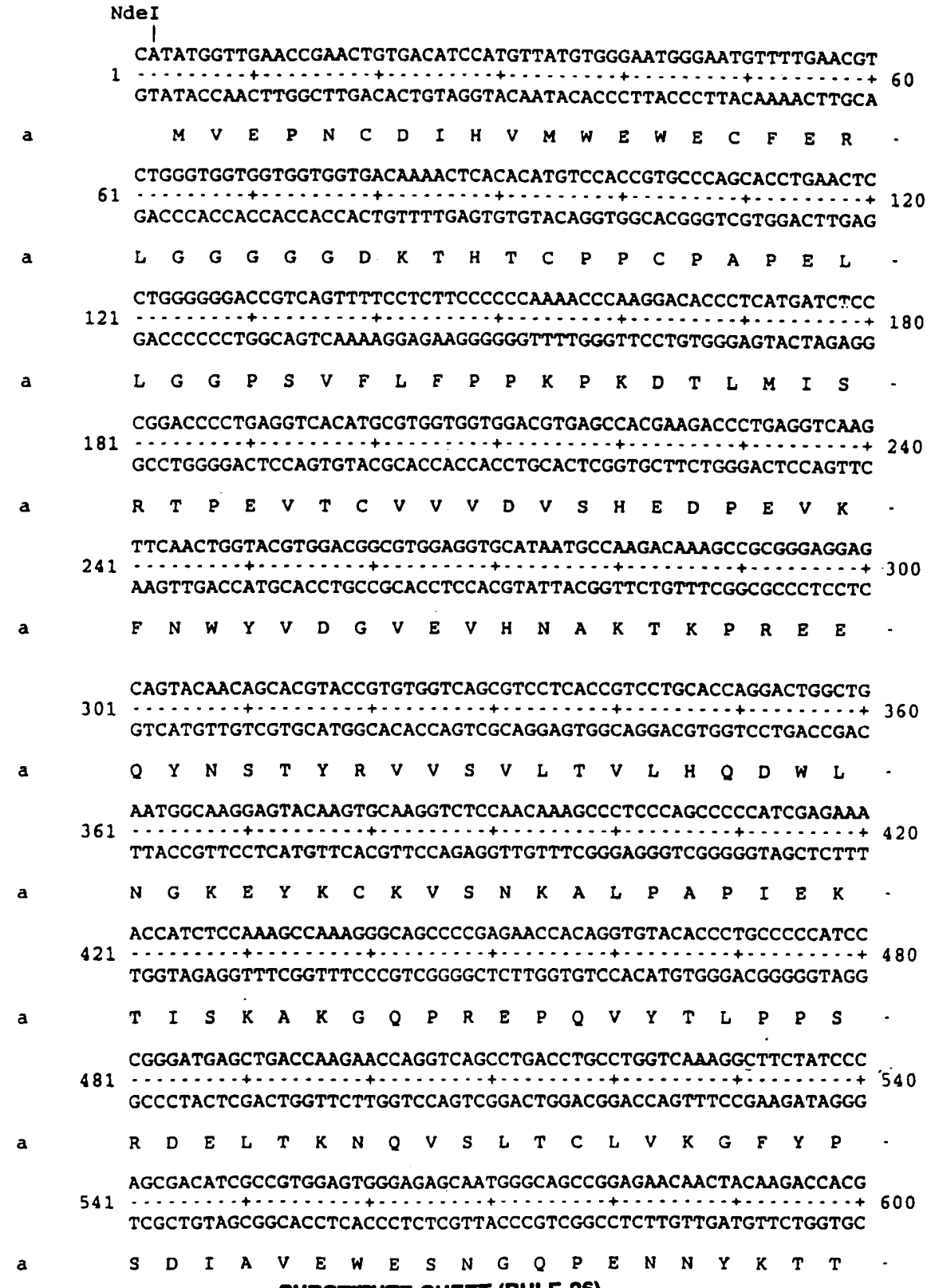

NdeI
CATATGGTTGAACCGAACTGTGACATCCATGTTATGTGGGAATGGGAATGTTTTGAACGT
GTATACCAACTTGGCTTGACACTGTAGGTACAATACACCCTTACCCTTACAAAACTTGCA
a M V E P N C D I H V M W E W E C F E R -
CTGGGTGGTGGTGGTGGTGACAAAACTCACACATGTCCACCGTGCCCAGCACCTGAACTC
GACCCACCACCACCACCACTGTTTTGAGTGTGTACAGGTGGCACGGGTCGTGGACTTGAG
a L G G G G G D K T H T C P P C P A P E L -
CTGGGGGGACCGTCAGTTTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCC
GACCCCCCTGGCAGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGG
a L G G P S V F L F P P K P K D T L M I S -
CGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAG
GCCTGGGGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTC
a R T P E V T C V V V D V S H E D P E V K -
TTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAG
AAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTC
a F N W Y V D G V E V H N A K T K P R E E -
CAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
GTCATGTTGTCGTGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGAC
a Q Y N S T Y R V V S V L T V L H Q D W L -
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAA
TTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTT
a N G K E Y K C K V S N K A L P A P I E K -
ACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCC
TGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGG
a T I S K A K G Q P R E P Q V Y T L P P S -
CGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCC
GCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGG
a R D E L T K N Q V S L T C L V K G F Y P -
AGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACG
TCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGC
a S D I A V E W E S N G Q P E N N Y K T T -

FIG. 24B

CCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAG
601
660
GGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTC
a P P V L D S D G S F F L Y S K L T V D K -
AGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAAC
661
720
TCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTG
a S R W Q Q G N V F S C S V M H E A L H N -
BamHI
CACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAACTCGAGGATCC
721
773
GTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTATTGAGCTCCTAGG
a H Y T Q K S L S L S P G K *

FIG. 25A

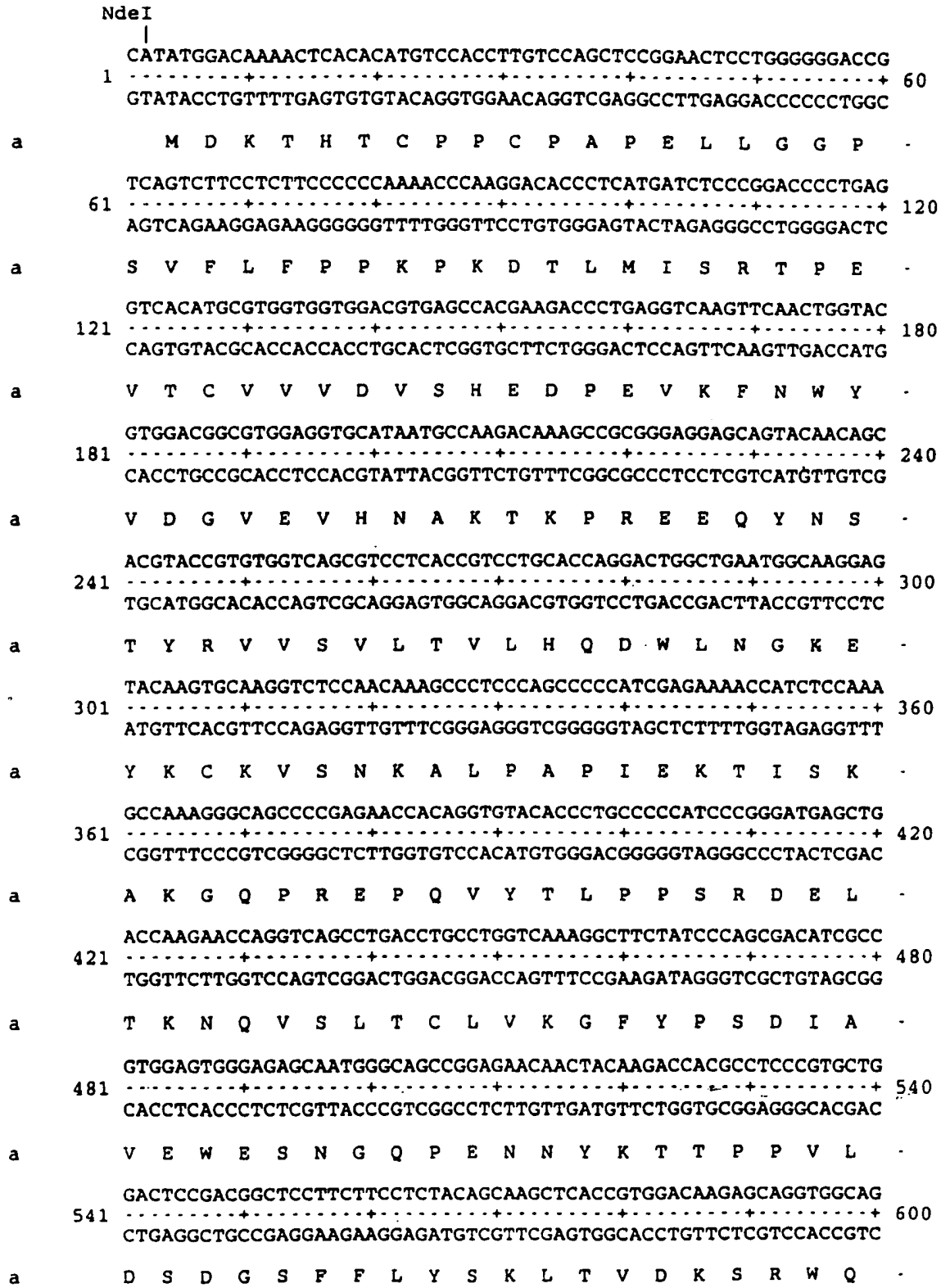

NdeI
CATATGGACAAAACTCACACATGTCCACCTTGTCCAGCTCCGGAACTCCTGGGGGGACCG
GTATACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGC
a M D K T H T C P P C P A P E L L G G P -
TCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
AGTCAGAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
a S V F L F P P K P K D T L M I S R T P E -
GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
a V T C V V V D V S H E D P E V K F N W Y -
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG
a V D G V E V H N A K T K P R E E Q Y N S -
ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC
a T Y R V V S V L T V L H Q D W L N G K E -
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
ATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
a Y K C K V S N K A L P A P I E K T I S K -
GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
a A K G Q P R E P Q V Y T L P P S R D E L -
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG
a T K N Q V S L T C L V K G F Y P S D I A -
GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTG
CACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGAC
a V E W E S N G Q P E N N Y K T T P P V L -
GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCCACCGTC
a D S D G S F F L Y S K L T V D K S R W Q -

FIG. 25B

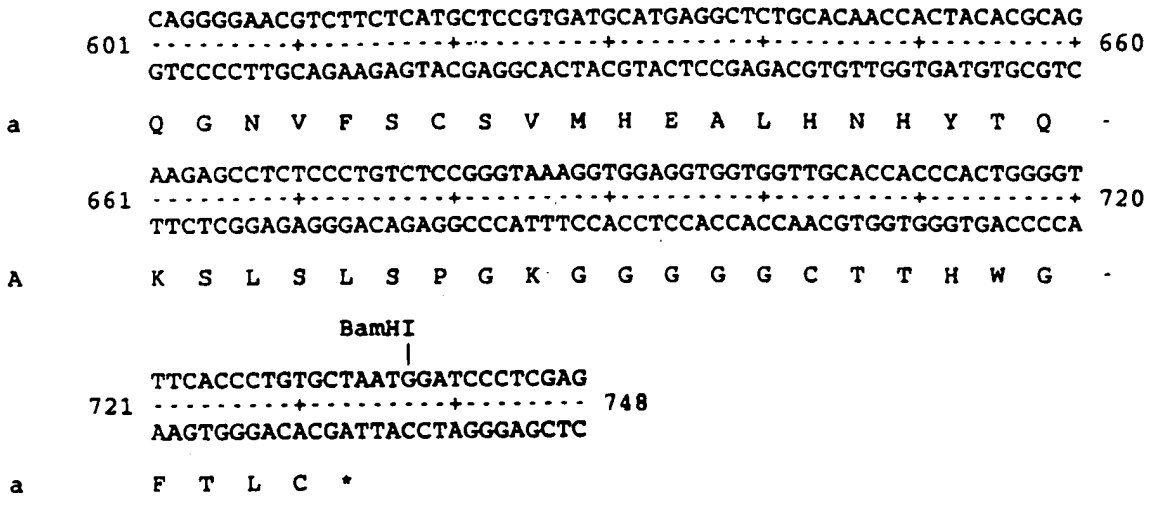

CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
601
660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a Q G N V F S C S V M H E A L H N H Y T Q -
AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTTGCACCACCCACTGGGGT
661
720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCAACGTGGTGGGTGACCCCA
A K S L S L S P G K G G G G G C T T H W G -
BamHI
TTCACCCTGTGCTAATGGATCCCTCGAG
721
748
AAGTGGGACACGATTACCTAGGGAGCTC
a F T L C *

FIG. 26A

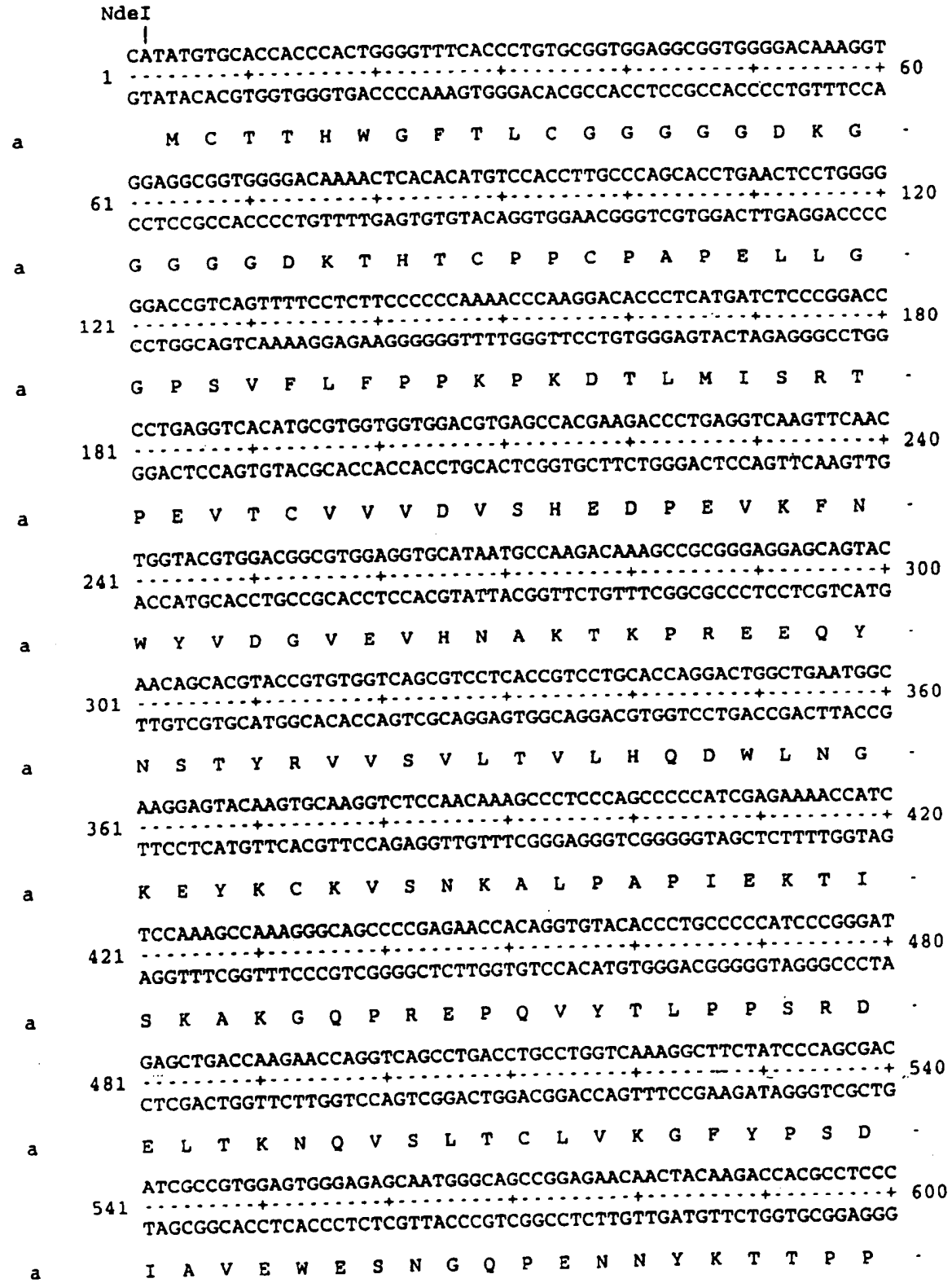
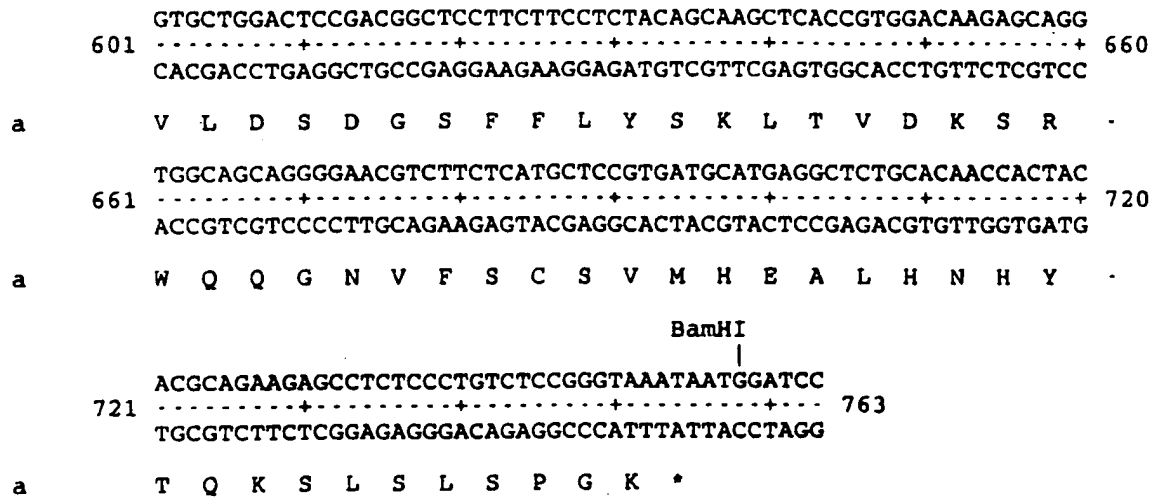

NdeI
CATATGTGCACCACCCACTGGGGTTTCACCCTGTGCGGTGGAGGCGGTGGGGACAAAGGT
GTATACACGTGGTGGGTGACCCCAAAGTGGGACACGCCACCTCCGCCACCCCTGTTTCCA
M C T T H W G F T L C G G G G G D K G
GGAGGCGGTGGGGACAAAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGG
CCTCCGCCACCCCTGTTTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCC
G G G G D K T H T C P P C P A P E L L G
GGACCGTCAGTTTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACC
CCTGGCAGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGG
G P S V F L F P P K P K D T L M I S R T
CCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAAC
GGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTG
P E V T C V V V D V S H E D P E V K F N
TGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTAC
ACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATG
W Y V D G V E V H N A K T K P R E E Q Y
AACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGC
TTGTCGTGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCG
N S T Y R V V S V L T V L H Q D W L N G
AAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATC
TTCCTCATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAG
K E Y K C K V S N K A L P A P I E K T I
TCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGAT
AGGTTTCGGTTTCCCGTCGGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTA
S K A K G Q P R E P Q V Y T L P P S R D
GAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGAC
CTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTG
E L T K N Q V S L T C L V K G F Y P S D
ATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC
TAGCGGCACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGG
I A V E W E S N G Q P E N N Y K T T P P

FIG. 26B


GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGG
601
660
CACGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTCGAGTGGCACCTGTTCTCGTCC
a V L D S D G S F F L Y S K L T V D K S R -
TGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTAC
661
720
ACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATG
a W Q Q G N V F S C S V M H E A L H N H Y -
BamHI
ACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCC
721
763
TGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGG
a T Q K S L S L S P G K *